COMPUTATIONAL AND CAUSAL APPROACHES ON SOCIAL MEDIA AND MULTIMODAL SENSING DATA: EXAMINING WELLBEING IN SITUATED CONTEXTS

A Dissertation
Presented to
The Academic Faculty

By

Koustuv Saha

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Interactive Computing
College of Computing

Georgia Institute of Technology

August 2021

COMPUTATIONAL AND CAUSAL APPROACHES ON SOCIAL MEDIA AND MULTIMODAL SENSING DATA: EXAMINING WELLBEING IN SITUATED CONTEXTS

Thesis committee:

Dr. Munmun De Choudhury,
Advisor
School of Interactive Computing
Georgia Institute of Technology

Dr. Gregory D. Abowd
College of Engineering
Northeastern University
Georgia Institute of Technology

Dr. Thomas Plötz
School of Interactive Computing
Georgia Institute of Technology

Dr. Emre Kıcıman
Information and Data Sciences
Microsoft Research

Dr. Gloria Mark
Donald Bren School of Information and Computer Sciences
University of California, Irvine

Date approved: June 3, 2021

Instinct is a marvelous thing. It can neither be explained nor ignored.

Agatha Christie

There's always some room for improvisation.

Satyajit Ray

To my Maa, who would have been the proudest and the happiest to see this day.

ACKNOWLEDGMENTS

I am very excited to write the acknowledgments section of this dissertation. As I write this section, I feel emotional and thankful to so many and many more, who have been instrumental in my Ph.D. journey and life in general. I am grateful to the guidance, support, and love I have received from my advisor, mentors, friends, family, and collaborators. My gratefulness and appreciations could barely be expressed in words, and is largely an injustice to what the above individuals deserve in making this dissertation possible.

First, I dedicate this dissertation to my mother, Sukla Saha, whom I could not ever thank directly. She dedicated more than half of her life to me, and it was only her sacrifices and relentless effort that has made me what and where I am today.

I am eternally indebted to my advisor, Dr. Munmun De Choudhury for shaping me as a researcher and as a person. I still remember how clueless I was about academic research when I joined the program. For the lack of a better explanation, I would say that my advisor *understood* my interests. My advisor enabled me to work on problems catering to the college students, a community that I was always personally motivated to and contribute towards. My advisor's advising has been the most optimal and balanced that suited my work style. She has not only guided and supported me at various stages and aspects of research but also enabled and encouraged me to pursue my research interests and skills. For instance, towards the end of the first year of my Ph.D., I expressed my interest in doing causal-inference research, which my advisor thoroughly supported, and it eventually became one of the key components of my research agenda. Time and again, I would want to study different problems, but my advisor's constructive criticism and feedback about the necessity of the problem would help me introspect into and shape the course of the problem. Besides, what I have learned from my advisor are not just limited to academics and research, she has helped me develop my perspectives and knowledge about daily and general aspects of life.

I am very fortunate to collaborate with and be mentored by some of the best and brightest

in the world. I have learned a lot from and be forever grateful to my committee members, Dr. Gregory Abowd, Dr. Thomas Plötz, Dr. Emre Kıcıman, and Dr. Gloria Mark. In all these years, I have had multiple meetings where I would literally shiver inside, worrying what if I could not answer a question. However, these interactions have significantly influenced me to think deeper and evolve as a researcher. I thank all the other PIs in the Tesserae project, Dr. Aaron Striegel, Dr. Andrew Campbell, Dr. Nitesh Chawla, Dr. Sidney D'Mello, and Dr. Anind Dey, all of whom, have played a pivotal part in making this dissertation happen. I thank Dr. Ingmar Weber, Dr. Michael Birnbaum, Dr. Louis Tay, and Dr. John Torous, who have helped me learn right from the early days of my Ph.D. I thank Dr. Amy Bruckman, Dr. Mark Riedl, Dr. Ashok Goel, Dr. Jacob Eisenstein, Dr. David Muchlinsky, and Dr. Rosa Arriaga for guiding and advising me in different stages of my Ph.D.

I am grateful to my amazing internship mentors, Dr. Naveen Ashish, Dr. Krishna Gummadi, Dr. Amit Sharma, and Dr. Maarten Bos. Besides research outcomes, my mentors helped to shape my research outlook and perspectives. I draw inspiration from my mentors to expand the depth and breadth of my thinking and approaches. I was exposed to new working styles, problems, and environments. More importantly, the internships helped me to form friendships to be cherished forever.

I do not have enough words to thank my friends, coauthors, and collaborators. They have not only been a network and pillar of support and wellbeing for me in all these years, but also I have looked up to them in uncountable aspects of life. This dissertation would never have materialized without my labmates, Dr. Stevie Chancellor, Sindhu Ernala, Vedant Das Swain, and Dong Whi Yoo. It was an honor to be a part of such an eclectic and bright group, who were both supportive as well as critical to drive the best out of me. I am grateful for the support-net you provided me during this bumpy Ph.D. ride. It is daunting to even think about the dissertation if I had not received the feedback and support from all of you in these years. I could share and seek help about literally everything to this group. I am going to miss the hangouts at Vedant's place and Sindhu's place, our "road-trips" and outings, and

the heart-to-heart conversations. I will miss the evening badminton sessions during the final few, and wish we had started these earlier. I acknowledge the work of Manikanta Reddy for contributing to this dissertation. I thank Mehrab Bin Morshed, Shagun Jhaver, Benjamin Sugar, Eshwar Chandrasekharan, Qiaosi Wang, Asra Yousuf, Pranshu Gupta, Jiawei Zhou, Sachin Pendse, Florina Dutt, Dr. Umashanthi Pavalanathan, Dr. Daejin Choi, Dr. Talayeh Aledavaood, Dr. Mai ElSherief, and so many others for being wonderful collaborators and friends to share the Ph.D. journey. I will thoroughly cherish the learning and enjoyable experiences I have had in academic and non-academic discussions. I am also thankful to the significant others of my closest friends, with whom I strongly bonded in the last few years, Sai Srirampur, Kate Dang, Eun Jeong Park, and Afifa Tasnim — we fought through the pandemic together. For a countless number of times, I have reached out to my friends for so many reasons, and I have never been disappointed. It has been such a pleasure to know and interact with you all, and I hope our paths keep crossing forever.

The CampusLife and Tesserae projects enabled me to form so many awesome partnerships and friendships, including with, Larry Chan, Dr. Kaya de Barbaro, Dr. Stephen Mattingly, Gonzalo Martinez, Ted Grover, Subigya Nepal, and Shayan Mirjafari; I thank you all a lot. I also thank my super cool internship friends, Vedant Nanda, Reza Babaei, Gaurav Parthasarathy, Ashish Kashinath, Venkatesh Kodukula, Ashwin Rajadesingan, Dr. Taisa Kushner, Nick Vincent, and Farhan Asif Chowdhury.

I thank my Tumlin roommates, Dr. Arkadeep Kumar, Dr. Dhwanil Shukla, and Dr. Anish Mukherjee to help me acclimatize with Atlanta. My friends from pre-Ph.D. life have showered love and care. I needed the frequent Skype calls with Vaibhav Pandey to re-fuel myself in the Ph.D. ride. I also doubt I would have done the Ph.D. had we (Vaibhav and I) not applied together. I thank my friends from Abzooba, Dr. Subhayan Mukerjee, Riti Chowdhury, Nirmalendu Biswas, Raj Nandy, Punit Kishore, Sayandeep Majumder, Sunit Sushil, Sumit Das, and Antarip Biswas — cheers for keeping the Whatsapp group alive, and keeping me updated about the gossips and happenings in our social circles! A shout-out to my friends from KGP, including Prateek Goyal, Lalit Yadav, Kuldeep Singh Chauhan, and Cherivirala Yaswanth Kumar. My friends from childhood, Sohini Mitra, Shweta Mohandas, Dr. Saurajyoti Kar — thanks for being in constant touch and helping me in several respects.

Special thanks to my family, my father, Dr. Surajit Saha, and my mother, Sukla Saha for their continued and colossal amount of support throughout my life. I thank Juhi Banerjee for sticking by my side in great and tough times, and for infusing confidence and encouragement in these years. I am fortunate to have them who never lost faith in me and were not hesitant to make sacrifices in order to prioritize my ambitions. Together, as a family, we will overcome the hardships and hurdles in the current and the future. I should also thank the technologies that have enabled us to be connected continuously and remotely, more so during the pandemic times, this time as a user (and not a researcher), — cheers to the Internet and its platforms like Whatsapp, Skype, Google, Facebook, Netflix, Amazon Prime, Hotstar, and so on. Ah, the family Whatsapp groups of *Sachin Leela* and *Great MMR* kept me up-to-date about the family affairs from being far away. I am thankful to the support and help from my extended family of uncles, aunts, and cousins in several regards. Speaking of internet technologies, I thank Youtube, Google Scholar, and Overleaf for their existence; concentrating and *writing* would have been much more difficult without you.

Finally, I acknowledge support from all at Georgia Tech in helping me navigate through the program. I acknowledge the contribution of every collaborator and coauthor who have made this work possible from so many aspects. The research included in this dissertation would not have been possible without the funding from IARPA, NIH, and NSF. I am honored to receive the support from GVU Foley Scholarship and Snap Research Fellowship to support my dissertation research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

A core aspect of our lives is often embedded in the communities we are situated in. The interconnectedness of our interactions and experiences intertwines our situated context with our wellbeing. A better understanding of wellbeing will help us devise proactive and tailored support strategies. However, existing methodologies to assess wellbeing suffer from limitations of scale and timeliness. These limitations are surmountable by social and ubiquitous technologies. Given its ubiquity and wide use, social media can be considered a "passive sensor" that can act as a complementary source of unobtrusive, real-time, and naturalistic data to infer wellbeing. This dissertation leverages social media in concert with multimodal sensing data, which facilitate analyzing dense and longitudinal behavior at scale. This work adopts machine learning, natural language, and causal inference analysis to infer the wellbeing of individuals and collectives, particularly in situated communities, such as college campuses and workplaces.

Before incorporating sensing modalities in practice, we need to account for confounds. One such confound that might impact behavior change is the phenomenon of "observer effect" — that individuals may deviate from their typical or otherwise normal behavior because of the awareness of being "monitored". I study this problem by leveraging the potential of longitudinal and historical behavioral data through social media. Focused on a multimodal sensing study, I conduct a causal study to measure observer effect in social media behavior and explain the observations through existing theory in psychology and social science. The findings provide recommendations to correcting biases due to observer effect in social media sensing for human behavior and wellbeing.

The novelties and contributions of this dissertation are four-fold. *First*, I use social media data that uniquely captures the behavior of situated communities. *Second*, I adopt theory-driven computational and causal methods to make conclusive research claims on

wellbeing dynamics. *Third*, I address major challenges with methods to combine social media with multimodal sensing data for a comprehensive understanding of human behavior. *Fourth*, I draw interpretations and explanations of online-data-driven offline inferences. This dissertation situates the findings in an interdisciplinary context, including psychology and social science, and bears implications from theoretical, practical, design, methodological, and ethical perspectives catering to various stakeholders, including researchers, practitioners, and policy-makers.

CHAPTER 1
INTRODUCTION

A core aspect of our social lives is often embedded in the communities that we are situated in, such as our workplaces, residential compounds, neighborhoods and localities, school and college campuses, or even physically co-located interest and demographic communities, including third places [448]. The inter-connectedness and inter-dependencies of our interactions, experiences, and concerns, make individual and collective wellbeing interlinked in situated communities. For example, an event of crime or violence in a neighborhood often causes alertness and anxiety among several neighborhood residents. A better understanding of psychosocial dynamics will also help devise strategies to address wellbeing concerns in situated communities. However, existing methods to assess wellbeing suffer from limitations of scale and timeliness. On the other hand, social media, for its ubiquity and widespread use, can be considered to be a "passive sensor" that can act as a complementary source of unobtrusive, real-time, and naturalistic data to infer wellbeing. Human behavior is a complex function of social, psychological, and environmental underpinnings which, when studied without minimizing confounding factors, may lead to unreliable and inconclusive findings. **By proposing computational and causal approaches that minimize the confounds, this dissertation leverages social media in concert with multimodal data to examine wellbeing in situated communities.** However, the feasibility of proactive and real-time social media technologies for wellbeing may be sensitive to further confounds in practice, which are invisible in research using retrospectively collected data. **This dissertation makes a case for one such concern, "observer effect" and examines its pervasiveness in social media behavior in multimodal sensing.** The subsequent paragraphs elaborate on the specific motivations and contributions of this dissertation.

Studies of human behavior and wellbeing have typically relied on self-reported survey data from individuals. These approaches suffer from a variety of limitations. For instance, self-reported data suffers from subjective assessments, recall and hindsight biases, and are typically retrospective— information is gathered after an event has occurred, or after an individual has experienced a specific change [620]. Recent research has recognized the value of in-the-moment data recording and acquisition approaches, such as via active sensing in the form of ecological momentary assessments (EMAs) that ask an individual to log their momentary state and activities [651]. However, active sensing is challenged with scale, access, and cost [565]. EMAs, which are often disseminated through prompts on smartphones, potentially induce response burden on participants [603]. This leads to a trade-off between balancing the construct validity of participant responses and compliance [109]. Subsequently, researchers have employed various forms of passive sensing [651], such as logging an individual's phone usage or tracking physical activity via wearable sensors, and these sensing technologies have successfully helped us study human behavior, wellbeing, and psychosocial dynamics [651].

This dissertation proposes social media as one such passive sensor. It provides an inexpensive and unobtrusive means to gather real-time and historical data of individuals in natural settings [529]. The premise is built on the findings of a growing body of work, which has leveraged social media to identify markers and to assess risk regarding a variety of psychosocial health and wellbeing concerns [166, 206, 280, 531]. Social media data captures people's linguistic expressions, therefore, a unique strength its ability to function as a *verbal sensor* to understand psychosocial dynamics. The potential of social media sensing for complex human behaviors and wellbeing is explained by the Social Ecological Model [102, 647]. This model implies that human behaviors and experiences are not isolated, and are impacted by our relationships, the communities that we are situated in, and by the events and factors in the society. **Therefore, to get a better understanding of wellbeing we need to include situated contexts, and this dissertation focuses on thinking about situated communities as examples to consider situated contexts.**

Situated communities are typically defined over geographic spaces where individuals share some form of physical colocation (same floor, same building, same locality, same campus, etc.). Here, individuals share social interactions and bear common and distinctive social ties and interests [486]. Members of situated communities often access common resources and institutions dedicated to the wellbeing and support [484]. From a wellbeing perspective, individual and collective wellbeing is closely associated in situated communities. Again, the absence of appropriate and proactive support strategies may exacerbate the overall wellbeing manifold owing to the inter-dependencies and inter-connectedness in situated communities. For instance, the lack of timely supportive interventions following an external crisis can proliferate community-cascading acute stress experiences leading to several negative consequences. An overwhelming amount of stress following a crisis can lead to long-term negative mental health outcomes, such as post-traumatic stress disorder, acute stress disorder, borderline personality disorder, or adjustment disorder [668].

However, it is challenging to capture the subjective aspects of individual lives in situated contexts [27]. As already noted, traditional and most existing approaches of understanding wellbeing suffer from limitations and are typically reactive in nature [620]. These approaches are largely based on discrete occurrences of events, and there is no way to continually and comprehensively assess wellbeing dynamics in situated communities.

This dissertation aims to overcome the gap of studying wellbeing in situated communities by using and complementing social media with multimodal data. In particular, this dissertation focuses on two common situated communities, which many of us can relate ourselves with, *college campuses* and *workplaces*. For example, within college campuses, I study the effect of gun violence events on stress levels of college students and the effectiveness of post-crisis interventions, particularly public service announcements on counseling recommendations following student deaths on college campuses. For workplaces, I study how collective workplace dynamics such as organizational culture or individualistic role dynamics influence individual wellbeing and performance at workplaces.

The **novelties and contributions** of this dissertation are four-fold: **First**, I use social media data that uniquely captures the behavior of situated communities. **Second**, I adopt theory-driven computational and causal methods to make conclusive research claims. **Third**, I address major challenges of social media by developing methods to combine social media with complementary multimodal sensing data for a comprehensive understanding of human behavior. **Fourth**, I introspect into drawing meaningful interpretations of online-data-driven offline inferences. This dissertation situates the findings in an interdisciplinary context, including psychology and social science, and bears implications from theoretical, practical, design, methodological, and ethical perspectives catering to various stakeholders, including researchers, practitioners, and policymakers. The next few paragraphs unpack the novelties and contributions of the dissertation.

This dissertation leverages social media data that reflect online analog of the offline and physically co-located situated communities. In the context of college subreddits, I leverage college subreddit data where college students express and share topics and interests about their day-to-day academic, personal, and college lives [530, 531, 544]. Similarly, in the context of workplaces, I leverage Glassdoor data where employees publicly express their workplace experiences [155], and LinkedIn data which is used as a professional social networking platform. These datasets uniquely capture individual disclosures amid the social and environmental context towards a better understanding of wellbeing.

We note that human behavior and wellbeing dynamics are influenced by several intrinsic and extrinsic factors in both normalcy and crisis. To make conclusive research claims from social media data, *I adopt causal-inference and computational approaches drawing on machine learning, natural language analysis, and statistical modeling*. Causal methods minimize the confounds and lead to stronger claims about cause-and-effect relationships in people's reactions to certain events or environments. For example, to understand the effect of gun-violence events on student stress, I minimize stress attributable to academic, personal, relationship, environmental, and other factors in college students' lives.

Despite its potentials, social media data comes with challenges, and *I propose methods to address some of the major challenges of social media data*, such as the lack of ground-truth and the lack of social media presence altogether. For this purpose, I augment social media with multimodal sensing data such as EMAs and passive sensing streams. Next, it is important to recognize that these computational assessments have potential real-world implications, so, we need to be careful, and meaningfully understand what we are measuring. *A cross-cutting theme in this dissertation is introspecting and interpreting online-data-driven offline inferences*. A significant component in this dissertation drawing insights and explaining the observations and potential consequences. For instance, this dissertation critically explores what life events people disclose on social media, and why?

Together, this dissertation makes methodological contributions in providing computational techniques and frameworks to measure wellbeing in situated communities, and makes the case for building rigorous but ethical approaches by critically reflecting on the practical and real-world consequences. I situate the findings in an interdisciplinary context including psychology and social science and bears implications from theoretical, practical, design, methodological, and ethical perspectives catering to a variety of stakeholders, including researchers, practitioners, administrators and policy-makers. A major implication concerns building tools that leverage these data-driven methodologies to improve wellbeing in practice. For instance, campus and workplace welfare staff can use these tools to continually track people's wellbeing and act proactively with timely and tailored interventions. However, in reality, prospective data collection and use may be different and bring new challenges, such as that of, "observer effect", or an individual's tendency to modulate and alter their behavior with the awareness of being observed [609]. If not accounted for, such a challenge would bring in new confounds and weaken the effectiveness of algorithms and approaches built on retrospective data. This dissertation provides a causal methodology to identify and measure observer effect in social media behavior and provides insights and recommendations on accounting for this behavior in prospective use of social media and multimodal sensing for human behavior and wellbeing.

Organization of the dissertation. The organization of the dissertation is as follows: Chapter 2 discusses the background and related work. Chapters 3 and 4 elucidates computational and causal methods of examining wellbeing with social media in situated contexts, where chapter 3 focuses on college campuses, and chapter 4 focuses on workplaces. Chapter 5 describes methods to overcome two challenges of social media as a "sensor" — a) lack of ground-truth and b) lack of social media data. Chapter 6 introspects into online-data-driven offline inferences to draw meaningful interpretations, particularly around the efficacy and tradeoffs of combining social media and offline sensing to personalize predictions, and the content and factors associated with social media disclosures, particularly those related to life events. Chapter 6 describes a challenge in prospective settings of social media based wellbeing sensing called as the "observer effect", and shows a study on measuring observer effect in social media behavior within the context of multimodal sensing.

Table 1.1: Summary of studies completed as a part of the dissertation.

Study	Thematic Area	Summary	Data	Location
Modeling Stress with Social Media Around Incidents of Gun Violence on College Campuses [531]	Social media study of wellbeing in situated communities (college campuses)	Building a machine learning classifier to infer stress in social media language. Use this classifier to study evolution of stress around gun violence incidents on college campuses by adopting a causal-inference based interrupted time-series approach.	Reddit	Chapter 3 (3.2)
A Social Media Based Examination of the Effects of Counseling Recommendations After Student Deaths on College Campuses [544]	Social media study of wellbeing in situated communities (college campuses)	This study adopted a causal-inference framework to examine the effects of counseling recommendations after student deaths on college campuses, and then measured affective, behavioral, and cognitive changes on social media after exposure. Another contribution of this study includes a grief lexicon in the Circumplex model of affect.	Reddit	Chapter 3 (3.3)
Modeling Organizational Culture with Workplace Experiences Shared on Glassdoor [155]	Social media study of wellbeing in situated communities (workplaces)	Adopting a theory-driven approach to model lexico-semantic word embedding representations of Glassdoor posts as organizational culture. Evaluate the our model of organizational culture explains workplace performance better.	Glassdoor, O*Net, Other Data: Surveys	Chapter 4 (4.2)
LibRA: On Linkedin based Role Ambiguity and its Relationship with Wellbeing and Job Performance [536]	Social media study of wellbeing in situated communities (workplaces)	Computational modeling role ambiguity as the lexico-semantic difference in LinkedIn description and job description. Examining that LibRA explains wellbeing and job performance better.	LinkedIn, O*Net, Other Data: Surveys	Chapter 4 (4.3)
Inferring mood instability on social media by leveraging ecological momentary assessments [529]	Addressing challenges of social media by complementary multimodal sensing (lack of ground-truth)	Leveraging limited ground-truth collected via EMAs to model mood instability on social media data. Adopting semi-supervised learning by using unlabeled and large-scale social media data to improve the classifier.	Facebook, Twitter, EMAs (Active Sensing)	Chapter 5 (5.1)
Imputing Missing Social Media Data Stream in Multisensor Studies of Human Behavior [537]	Addressing challenges of social media by complementary multimodal sensing (missing social media data)	Leveraging physical sensor behavior to predict latent dimensions of (missing) social media behavior. Evaluating the performance of the imputation framework in predicting individual differences in psychological traits.	Facebook (Social Media), Wearable, Bluetooth, Smartphone (Passive Sensing), Surveys	Chapter 5 (5.2)
Contextualizing Person-Centered Predictions with Social Media [532]	Introspecting into online-data-driven offline inferences	Building person-centric models of predicting psychological contructs through social media by contextualizing on offline behaviors as captured by multimodal passive sensing. Examining the efficacy and tradeoffs of personalization efforts.	Facebook, Multimodal Passive Sensing, Surveys	Chapter 6 (6.1)
Understanding Life Event Disclosures on Social Media [538]	Introspecting into online-data-driven offline inferences	Examining what life events are disclosed on social media, and what event-centric and individual-centric factors explain life event disclosures on social media	Facebook (Social Media), Surveys, PERI life events survey [181]	Chapter 6 (6.2)
Measuring Observer Effect in Social Media Behavior	Observer Effect in Social Media Behavior	Examining if people alter social media behavior in prospective data collection settings. Adopting time-series and causal-inference analysis methods to measure the deviation in actual behavior from expected behavior post-enrollment in a multimodal sensing study. Explaining how this behavior varies with intrinsic traits.	Facebook (Social Media), Surveys	Chapter 7

CHAPTER 2
BACKGROUND

2.1 Situated Communities

"Situated communities consist of geographically co-located, diverse, and close-knit communities where individuals share distinctive social ties" (Bin Morshed et al., 2019) [63, 486, 544]. Importantly, spatial and contextual attributes influence the wellbeing of individuals and collectives within situated communities [233]. Therefore, ensuring that the members cope with psychological and cognitive demands is essential for both individual and collective wellbeing. This requires identifying and understanding psychological changes in circumstances of both normalcy and crisis. According to Murphey (1999), defining a community should be grounded in locally meaningful realities [434]. This definition has been adopted to study several forms of situated communities varying in sizes, granularity, and definitions, which include neighborhood, school districts, urban or rural areas, college campuses, and so on [484]. This dissertation focuses on problems that concern the wellbeing in two kinds of situated communities, *college campuses* and *workplaces*. Both these types of communities are unique in age, demographics, and socio-economic characteristics, as well as day-to-day activities, goals, and concerns.

2.1.1 Wellbeing in College Campuses

According to 2018 AUCCD survey, the most frequent concern for college counseling centers around the world are anxiety (58.9%), followed by depression (48.0%), stress (46.9%), specific relationship problems (29.5%), family concerns (29.0%), suicidal thoughts (28.4%), academic performance difficulties (28.2%), sleep disturbance (19.1%), social isolation or loneliness (18.5%), significant previous mental health treatment history (16.5%), and adjustment to a new environment (15.8%) [373]. Although these numbers are already a significant proportion, mental health concerns are known to be under-reported, more so for college students [195]. In particular, these numbers are largely based on what is reported or what the student seek support and care for, and given the stigma surrounding mental health concerns, actual proportions are plausibly higher. Eisenberg and colleagues have extensively studied the mental health problems and major impediments of seeking mental health care among college students [195, 196]. Bayram and Bilgel has explored the prevalence and socio-economic correlations of various mental health concerns among college students [50]. Literature has also outlined major initiatives and factors specifically concerning intervention initiatives following crises on college campus [66, 432].

Among these concerns, stress constitutes one of the most significant and prevalent concerns. As much as three out of four college students consider themselves to be stressed [379]. A variety of personal and academic life factors and environmental stressors precipitate college student stress [518]. Stressful episodes, in turn, are associated with cognitive deficits in students (e.g. concentration difficulties), decreased life satisfaction, and poor health behaviors [677]. Due to these multi-faceted risk factors and consequences of the stress experience, a variety of techniques have been employed to devise global and event-specific measures of stress in college students. For example, several investigations have modified life-event scales in an attempt to measure global perceived stress [128]. However, due to reliance on a specific list of events, this approach is insensitive to stress emanating from unforeseen or unanticipated circumstances such as crises. Additionally, subjective measures of response to specific stressors have been devised [247]. However, it can be difficult and time-consuming to adequately develop and validate an individual measure every time a new stressor is identified, such as following an environmental upheaval [130]. The Perceived Stress Scale [130] addresses many of these challenges in stress measurement and has been employed in studying college student stress [518].

However, data obtained from psychometric instruments are sensitive to self-report and retrospective recall biases, and may not necessarily reveal factors associated with the stigma of stress. Moreover, such measurements can only be conducted periodically, posing difficulties in understanding the temporal evolution of college student wellbeing. To address these limitations, recently, researchers have employed wearable sensing technologies and experience sampling methodologies to obtain real-time information on psychological symptoms [63, 239, 311, 651]. Although these techniques capture rich and dense behavioral signals, these kind of data collection suffers from limitations related to scale and compliance as they seek to actively engage the participants. In a parallel thread of research, psycholinguistics has revealed that language can serve as an indicator of wellbeing (stress, anxiety, depression), e.g., essays written by college students [474, 523, 659]. Drawing motivation from these these lines of research, this dissertation explores social media as a passive and verbal sensor to assess wellbeing of college students, particularly surrounding crises on college campuses, such as gun violence [531] and student deaths [544].

2.1.2 Wellbeing in Workplaces

Employee job satisfaction is of prime interest to both individuals as well as organizations. The complexities related to an individual's job role, or the *expectations applied to an individual within and beyond an organization's boundaries* can impact their job satisfaction [639]. The wellbeing of individuals at workplace also translates to individual, collective, and organizational success [350]. A rich body of literature in organizational studies, organizational psychology, and organizational behaviors has extensively studied the causes and correlates of improving wellbeing and performance at workplaces [61, 84]. Research postulates that wellbeing in workplace is an outcome of the interaction between individual characteristics and those of the working and organizational environment [61].

It is important to emphasize and address the challenges of workplace stress– *stress that arises if the demands of an individual's roles and responsibilities exceed their capacity and capability to cope*. Work-related stress is the second most compensated illness in Australia, and in the U.S. businesses face losses of over $30 billion a year owing to work-related stress. In order to understand workplace stressors better, De Neve et al. (2013) proposed the importance of considering subjective wellbeing at workplaces as a coarse construct that leads to objective benefits across the major life domains of 1) health and longevity, 2) income, productivity, and organizational benefits, and individual and social behavior [170].

Scholars note that survey questions may not only be responded "carelessly" [339], but also, because of their underlying parsimonious design, may not be interpreted the same way across individuals and groups, which may bias survey response. More pertinently, the use of survey methods for assessing worker wellbeing at the population-level is limited. It is costly to implement surveys at scale to assess worker wellbeing with high temporal granularity. Further, Pew survey estimates that response rates for surveys are very low (9% within the U.S.) [33, 45, 631] and will get lower as time passes due to the use of caller ID and spam protection. This leads to concerns of representativeness at large. This dissertation explores a new form of data to address the issue of understanding worker wellbeing by capturing people's naturalistic, self-motivated, and self-initiated expressions on social media to infer workplace wellbeing. By grounding the online-data driven offline inferences in a theory-driven approach, and validating with other gold-standard existing approaches, this dissertation illustrates the potential of social media and complementary multimodal sensing as a proactive tool to inform interventions for improving workplace wellbeing.

2.2 Social Media as a Passive Sensor

Research has demonstrated that social media technologies have a number of benefits as a passive sensing modality. In particular, it is low-cost, large-scale, non-intrusive to collect, and has the potential to comprehensively reveal naturalistic patterns of mood, behavior, cognition, psychological states and social milieu, both in real-time and across longitudinal time [252]. Considerable research has focused on developing approaches that can (semi-) automatically assess health and wellbeing by employing social media as a "sensor" for both individual- and population- centric assessments [147, 160, 166, 203, 277, 542, 543]. These studies reveal that social media technologies provide a number of benefits as a passive sensing modality [540].

2.2.1 Social Media for Predicting and Understanding Wellbeing

Literature in psychology has revealed that language can help us understand the psychological states of an individual [473]. In recent years, several studies have demonstrated that social media data can be analyzed to reliably infer and understand the psychological and mental health states of individuals and communities [592]. Research has leveraged social media data at scale to quantitatively identify conditions and symptoms related to diseases [468], disease contagion [526], mood and depressive disorders [166], mental health [65, 137, 540], post-traumatic stress disorder [138], eating disorders [111], suicidal ideation [167], psychotic symptoms [206], addictive behaviors [428], grief [82, 241], and substance and drug use [112, 540]. From the standpoint of collective wellbeing, Culotta et al. inferred county-level mental health using Twitter data [147].

Relatedly, social media has facilitated analyzing personality traits and their relationship to psychological and psychosocial wellbeing, through machine learning and linguistic analysis [356, 494, 564]. In parallel, crisis literature has also found promising evidence of supporting the potential of web and social media language to better understand the psychological impacts of external events and crisis [168, 397, 460, 597]. This body of work shows that online platforms emerge as a safe haven for people, enabling them to interact and express themselves during times of upheavals in their environment [23, 220, 397, 597]. Notably, Cohn et al. studied psychological markers in social media language post the 9/11 disaster [131]. The potential of social media in predicting individual and collective wellbeing is situated in the Social Ecological Model (see Figure 2.1)

Figure 2.1: Social Ecological Model: Human behaviors are deeply embedded in the complex interplay between an individual, their relationships, their communities, and societal factors. Social media provides a passive way to gather quantifiable signals about the social ecological dimensions relating to an individual's behavior [102, 149].

2.2.2 Social Media and Wellbeing in Situated Communities

Social Media and Wellbeing in College Campuses

Pertaining to the population of college students, Ellison et al. (2007) in a seminal work, found positive relationship between social media usage and maintenance of social capital [200] and Manago, Taylor, and Greenfield found that social media helps college students to satisfy enduring psychosocial needs [392]. Given the ubiquity of social media use among youth [479], and because social media platforms enable individuals to share and disclose mental health issues [197], researchers have begun to leverage social media as an unobtrusive and passive source of data to infer and understand mental health and wellbeing of college students [381, 399, 429].

Of particular relevance is Bagroy et al.'s work who built a collective mental health index of colleges employing social media (Reddit) data [31]. Manago et al. found that social networking helps in satisfying psychosocial needs of college students [392], and Moreno et al. studied mental health disclosures by college students on social media [429]. Prior research has also inferred other behavioral attributes and psychological attributes of college students, using social media [399, 630]. Recently, Saha, Yousuf, Boyd, Pennebaker, and De Choudhury have demonstrated the construct validity of assessing mental health on college students' social media data with respect to mental health consultations on college campuses [545].

Social Media and Wellbeing in Workplaces

In the last decade, researchers have used social media technologies to understand employee behavior [159]. In a seminal work, Ehrlich and Shami compared employees' use of social media platforms, particularly their motivations in their use of social media (Twitter) [193]. This work reports that social media use (both at home and work) made workers, especially mobile workers, feel more connected to other employees, and provided an avenue to boost personal reputation at the workplace. Studies have found that social media use is positively correlated workplace wellbeing [573]. Increased social media interactions within the workplace, through platforms such as IBM's Beehive, have been found to improve personal and professional networking, career advancements, and innovation [177, 178, 210, 235]. Other works find positive relationships between workplace and employee behavior, such as wellbeing, experiences, and engagement through social media technologies [159, 202, 224, 275]. In an early work, Skeels and Grudin conducted a longitudinal study of the motivations and use of social media platforms by workplace employees [582].

Again, social media and online engagement platforms have facilitated an effective means to study employee behavior and satisfaction — a body of research that is extensive in CSCW and HCI area [25, 159, 425, 572, 574, 582]. A variety of analytical and computational approaches on language and network dynamics have been applied to glean correlates of employee job satisfaction and wellbeing, such as engagement [299, 425, 572], employee affect [159, 537], social pulse [573], reputation [322], organizational relationships [85, 237, 424], workplace behavior [398], and job satisfaction [546]. Anonymized platforms like Glassdoor provide "safe spaces" for employees to share and assess workplace experience [74, 353]. Glassdoor data was used to model brand personalities based on *employee imagery* factors such as working conditions, company culture, and management style [674]. Lee and Kang used Glassdoor data to study the influence job satisfaction factors, and their influence on employee retention and turnover [371]. These studies indicate the value of such unobtrusive data sources in understanding workplace experiences.

In the professional networking space, LinkedIn has emerged as the primary social media platform [606, 628]. This platform, which was initially viewed as a "repository of webpages", gradually evolved to be informally known as "Facebook in a suit" [633], which serves as an online social space enabling individuals to enhance professional visibility [25, 95, 354]. LinkedIn allows the individuals to self-describe and self-promote their professional portfolio to either seek for new jobs, or to use it as their professional networking and webpage. Guillory and Hancock found that the public-facing nature of LinkedIn influences an individual's accountability and reduces deception in their self-description of their professional portfolio, which also aligns with Donath's early research on identity and deception in online spaces [183]. Researchers have studied the differences and similarities in the self-presentation behavior and use of LinkedIn in comparison to personal social media platforms such as Facebook and Twitter [25, 582, 633, 682]. Also, organizations' use of LinkedIn has grown tremendously over the years, which also implicitly puts peer- and societal- pressure on individuals to own and maintain LinkedIn accounts [351]. Utz and Breuer recently studied the individual-specific factors that influence their behavior on LinkedIn in terms of networking and informational benefits that the platform facilitates [629], Van et al. inferred personality traits on LinkedIn self-presentations of individuals [642], and Zide et al. studied how LinkedIn profiles differ across occupations [685]. Zhang, De Choudhury, and Grudin studied employees' privacy perceptions on social media [682].

While all these sources fall under a broad umbrella of "social media", the motivations to use any of these platforms might differ based on individual and platform-specific characteristics [25, 159, 193, 531, 682]. These factors are associated with both opportunities and challenges in leveraging social media data to understand workplace dynamics. This dissertation draws motivation from the above body of research to leverage Glassdoor data to measure organizational culture (Chapter 3.1) and LinkedIn data to measure role ambiguity

(Chapter 3.2). These chapters also show how these online-data-driven metrics are associated with the individuals' job performance and wellbeing.

2.3 Social Media in Concert with Multimodal Sensing

With the ubiquity of smartphones and wearables, passive sensing modalities enable convenient means to obtain dense and longitudinal behavioral data at scale [651, 653]. However, such a data collection is prospective — after enrollment, during the study period. To obtain historical or before-study data, researchers have recently used social media as a passive sensor, which enables unobtrusive data collection of longitudinal and historical data of individuals that were self-recorded [529, 531].

Together, passive sensing modalities in conjunction propagate the vision of "people-centric sensing" [96], although each one of them may have its own limitation. In the case of social media, it suffers from data sparsity, and not everyone is *equally active* on it [529, 680]. Therefore, the variability in the use of social media across individuals may impact the predictive capabilities of models built on "all" individuals' datasets — e.g., some features may have high-variance, some features' effects may be washed out, and some features may be downplayed by other features, although these could bear significant signals for certain cohorts of individuals. Again, it can function as a "sensor" only on those who use it. This leads to a common problem that many multimodal sensing studies of human behavior face [370, 529, 651]— they either examine a larger pool of participants with fewer sensors, or a smaller pool of participants who comply with all sensing streams. This compromises the combined potential of multiple sensors or the wide spectrum of individual behaviors.

The above literature motivates this dissertation in combining the complementary strengths of multimodal sensing through computational approaches to infer latent behavioral attributes [62, 154, 179, 472, 497]. Further, I develop methodologies to contextualize predictions on social media data per cohorts of "similar" individuals on the basis of their offline and physical behavior (captured using in-the-wild sensing technologies).

CHAPTER 3

SOCIAL MEDIA STUDIES OF WELLBEING ON COLLEGE CAMPUSES

College campuses are close-knit, largely geographically co-located communities, where students undergo several concerns related to mental wellbeing [196]. Colleges are valued institutions that help build upon a society's foundations and serve as an arena where the growth and stability of future generations begin. College students undergo stress throughout the year due to academic, personal, relationships, environmental, and social factors [518]. The pervasiveness of stress, depression, anxiety, hopelessness, and suicidal behavior is significant among college student population [302]. Besides, crises on college campuses can cause acute stressful experiences and can exacerbate into long-term negative consequences such as post-traumatic stress disorder, acute stress disorder, borderline personality disorder, or adjustment disorder [668].

However, methods to assess stress experiences or emotional responses of college students are plagued by challenges in access to timely information, exacerbated by the social stigma of the condition, lack of awareness of the condition, and the noted acceptance of stress in colleges as a "badge of honor" [31]. Further, campus mental health services services often lack in resources, staff, and preparedness, leading to long waiting lists and selective/infrequent consultations of many [232]. This understates an urgent need to meet the rising demand of mental health services with adequate and accessible resources. Currently, campus mental health services do not have adequate means to assess the evolving nature of demand or needs. While periodic surveys of students' mental health provides some barometer of mental health incidence, in terms of medication use, daily lifestyle, suicidal thoughts, depression symptoms, as well as potentially contributing academic, environmental, personal, and social factors [50], they are accurate only in snapshots, and are prone to retrospective and susceptible to biases, and may miss out real-time, fine-grained information or sudden fluctuations in psychological signals [620]. Self-reported and active solicitation techniques to students' stress vulnerabilities also suffer from social desirability and stigma associated with stress, and are often under-reported [31].

To overcome such limitations, passive sources of data have recently been explored, which provide dense and longitudinal data at scale [651]. Given the ubiquity and widespread use of social media, especially among the college student demographic, social media data has also been leveraged as a "passive sensor" that can act as a complementary source of unobtrusive, real-time, and naturalistic data to infer wellbeing [529]. Social media data is low-cost, large-scale, non-intrusive to collect, and has the potential to comprehensively reveal naturalistic patterns of mood, behavior, cognition, psychological states and social milieu, both in real-time and across longitudinal time for individuals and collectives [252]. Social media language consists of an individual's personal and social discourse about day-to-day concerns, and effectively reflects their health and psychosocial wellbeing in a variety of states and contexts [75, 194, 324]. Linguistic cues and social interactions on social media platforms have therefore, enabled researchers to study psychopathologies including depression, anxiety, stress, and loneliness [136, 166, 279, 540, 564].

Along the same lines, this chapter illustrates methods of examining wellbeing on college campuses through social media. In particular, I show two studies which are in the context of crises on college campuses, one is around gun violence incidents on college campuses, and another is around public service announcements (PSAs) of counseling recommendations sent out after student deaths on college campuses. For both of these studies with social media data from online college communities on Reddit.

The first study examines how college student stress evolves around gun violence incidents on college campuses. I first build a machine learning classifier of stressful expressions in social media language. Next, focusing on 12 incidents of campus gun violence between 2012-2016, and social media data gathered from college Reddit communities, this study reveals amplified stress levels following these incidents, which deviate from usual stress patterns on college campuses. I examine the linguistic changes around these gun violence incidents to find distinctive characteristics, such as decreased cognition and academic career-related conversations, but increased self-attention, social orientation, death and family-related conversations, and the emergence of collective identity and solidarity.

The second study examines the effectiveness of post-crisis intervention measures in the form of counseling recommendations after student deaths on college campuses by adopting a causal inference framework on social media data. I employ statistical modeling and natural language analysis to measure the psychosocial shifts in behavioral, cognitive, and affective expression of grief in individuals who are "exposed" to the counseling recommendations, compared to that in a matched control cohort. Drawing on crisis and psychology research, the findings suggest that individuals exposed to counseling recommendations show greater grief, psycholinguistic, and social expressiveness, providing evidence of a healing response to crisis and thereby positive psychological effects of the counseling recommendations.

3.1 Social Media Data of College Communities

We begin by explaining the data of the two studies on measuring wellbeing on college campuses. I use social media data of college communities from *Reddit*.

3.1.1 Why Reddit?

According to a recent Pew Research survey [479], over 90% of U.S. youth use social media. Reddit is one of the most popular social media platforms which caters to the age group between 18-29 years: 65% of Reddit users are young adults [479]. This age demography aligns well with the typical college student population, making Reddit a suitable choice for studying college communities.

Reddit is a social discussion website consisting of diverse communities known as "subreddits" that offer demographic, topical, or interest-specific discussion boards. Many colleges have a dedicated subreddit community, which provides a common portal for the students on campus to share and discuss about a variety of issues related to their personal, social, and academic life. The college subreddits name themselves after the college communities which they represent and often customize their pages with college logos and campus images to signal their identity. The subreddit pages use personalized Reddit icons and member names based on college nicknames and mascots. Table 3.1 shows examples of college subreddits, along with self-descriptions, number of members (and moderators) and personalized titles and icons.

The above observations indicate that college communities on Reddit are a reasonable data source to study the research questions in this dissertation. Bagroy, Kumaraguru, and De Choudhury established that Reddit data of college communities could be used to estimate campus-wise mental wellbeing. This study also showed that college subreddit data adequately represents the rough demographic distribution of the campus population of over 100 U.S. colleges, is sufficiently widely adopted in these college campuses, and can be employed as a reliable data source to infer the broader college communities' mental wellbeing [31]. Recent research revealed that predicting mental health from college subreddits bears construct validity with mental health consultations on college campuses [545]. While college students likely use other social media platforms as well, such as Facebook, Twitter, Instagram, and Snapchat, obtaining college-specific data from these sources is challenging because many of these platforms restrict public access of data, and they lack defined community structures. This induces difficulty in identifying college students and their college-related discussions, unless they self-identify themselves, which can limit both scalability and generalizability.

3.1.2 Collecting College Subreddit Data

The first step of data collection includes identifying college-wise subreddit data. For this, I take help of the SnoopSnoo website [586], which groups subreddits into several categories, one of which is "Universitites and Colleges". Within the United States, 174 out of the top 200 major ranked colleges [438] have a subreddit community, at an average of 3000

Table 3.1: Top 16 college subreddits by member count (June 2018). These subreddits often customize their page icons and member names with college nicknames or mascots.

Subreddit	Self-Description	#Members	Icon
r/UIUC	This subreddit is for anyone/anything related to UIUC	18,900 Illini	
r/berkeley	GO BEARS!	14,280 bears	
r/aggies	Anything Texas A&M community related!	13,477 Ags	
r/gatech	A subreddit for my dear Georgia Tech Yellow Jackets.	13,295 readers	
r/UTAustin	Welcome to The University of Texas at Austin	13,101 Longhorns	
r/OSU	The Ohio State University	13,348 4-string QBs	
r/ucf	Reddit for University of Central Florida	13,348 Knights	
r/UCSD	Members associated with the UC San Diego.	9,680 Tritons	
r/rutgers	For news relevant to Rutgers University.	9,654 Scarlet Knights	
r/VirginiaTech	A reddit for Hokies	9,605Hokies	
r/Purdue	Purdue University's subreddit.	9,602 Boilermakers	
r/rit	Rochester Institute of Technology official subreddit.	9,304 RITedditors	
r/UMD	The official subreddit of the University of Maryland.	9,172 ReddiTerps	
r/uofm	University of Michigan subreddit	8,756 wolverines	
r/ucla	A place for UCLA students, faculty, and fans! Go Bruins!	8,912 Bruins	
r/ASU	Subreddit for Arizona State University	8,316 Sun Devils	

Table 3.2: Example paraphrased comments in the college subreddit dataset.

HBD! I'd avoid beer if you've never drank for your first drink.
Join sublease FB group. You're probably looking for campus lodge, as they're cheap and allow pets.
Also, I was a bit confused [..] is Math 220 r ISAT 251 apart or is it a pre-requirement to get in?
Definitely get security package. I didnt on my bike and it got stolen last year [..]
im taking 498. its easy, but we did 16 chapters in 12 weeks whole semester its group presentations.
I am sorry, first of all. That's a loss that nobody should ever suffer, and certainly not during school.
Don't have sex with anyone on your floor unless you are committed to a year-long relationship.
Don't go partying instead of orientation week. I did and I regret it to this day.
You will get stressed and you'll have to learn how to deal with it.
I really don't want our school to get some sort of reputation, I hope whoever was involved is okay.
We have a good English dept, yet no journalism program. This college does sound great.

members each, and the largest college subreddits like *r/UIUC*, *r/berkeley*, *r/aggies*, *r/gatech*, *r/UTAustin*, *r/OSU*, and *r/ucf* have between 13K and 19K members.

To collect subreddit data, I access public archives of Reddit dataset hosted on Google BigQuery [256]. BigQuery is a cloud based managed data warehouse that allows third parties to access large publicly available dataset through SQL-like queries. This allows us to collect both posts and comments in the subreddit along with their author usernames and date timestamps, besides other metadata information (such as number of upvotes/downvotes). Table 3.2 reports a random sample of 10 comments from college dataset. A quick manual inspection of the comments on college subreddits suggests that these are about both seeking and sharing information and opinion on a variety of topics spanning across greetings, dorms, academics, partying, food, leisure, relationship, emotional support, and other miscellaneous aspects of college life in particular and youth life in general.

3.2 Examining Stress Around Gun Violence Incidents on College Campuses

Stress is a psychological reaction that occurs when an individual perceives that environmental demands exceed his or her adaptive capacity [570]. One of the populations particularly vulnerable to stress is college students [50]. Stress is identified as one of the major impediment to academic performance and student retention in colleges [677]. When stress becomes excessive, students experience physical and psychological impairment, and intensified stress can undermine resilience factors, such as hope.

However, external factors are also known to exacerbate college student stress [518]. A prominent set of such environmental attributes includes exposure to traumatic and violent events, which can profoundly impact college students' perceived stress and stress responses [560]. Overwhelming amounts of stress from such acute exposure can affect students' ability to cope, and regulate their emotions. Persistent stress episodes may bear long-term negative consequences [668].

Violent incidents on college campuses, ranging from mass shootings to acts of terrorism have proliferated in the recent past. A survey from Everytown for Gun Safety Support Research[1] reports that between 2013 and 2016, 76 incidents of gun violence have occurred on U.S. college campuses, resulting in more than 100 casualties. Many of these incidents not only affect those involved in the incidents directly, but also leave profound negative psychological impacts on the general campus community [677]. It is vital to understand the impacts that violent incidents have within the psyches of college students.

To reiterate, measuring wellbeing and psychological reactions in a proactive fashion is not only difficult given the challenges of existing methodologies, but also violent incidents

[1] everytownresearch.org Accessed: 2017-04-09

may further amplify these limitations. Due to the unique circumstances presented to a community exposed to a crisis, rehabilitation efforts that use student contributed data on stress and well-being are likely to be difficult to implement in a timely fashion [620].

This study addresses the above gaps by measuring the perception of stress on college campuses around gun violence, as manifested on social media. This work is motivated from two complementary research directions. Studies in psycholinguistics and crisis informatics have found promising evidence that the language shared on social media can help us infer psychological states of individuals and collectives [168, 397, 597]. Also, over 90% of young adults, or individuals of college going age use social media [262], providing a promising opportunity to study college students' mental wellbeing passively using social media data [31]. This study focuses on the following three research questions:

RQ1. How to automatically infer the stress expressions in social media posts?

RQ2. How stress expressions temporally change following gun violence events on college campuses?

RQ3. How stress expressions linguistically change following gun violence events on college campuses?

I focus on 12 gun violence incidents reported on U.S. college campuses between 2012-2016. For each campus, I collect data from their subreddit communities. Targeting RQ1, I develop an inductive transfer learning approach to infer stress expressed in Reddit posts, which achieves a mean accuracy of 82%. Using this classifier, I identify high stress posts shared in the 12 college subreddits. Then, targeting RQ2 and RQ3, I develop computational techniques drawing from time series and natural language analysis to assess the extent to which expressions of high stress change in the aftermath of the gun violence incidents.

3.2.1 Data and Methods

Gathering Campus-Specific Gun Violence Data

This study adopts the definition of gun violence on college campuses as published by Everytown for Gun Safety Support Research[1] – *"a shooting involving discharge of a firearm inside a college building or on campus grounds and not in self-defense"*. Everytown for Gun Safety is an American nonprofit organization, which conducts gun violence research in the U.S. Given the lack of a single database for gun violence incidents on college campuses, I adopt a snowball approach to curate our dataset [28, 469] – 1) I collect a seed list of gun violence incidents on US college campuses from Everytown for Gun Safety Research; also used in prior work [28]. 2) I augment this seed list with additional incidents that qualify the same definition as above — I consult credible online sources in an iterative fashion[2].

The curated list consist of gun-related violence incidents in and within a close proximity of a US college campus, all that happened between 2012 and 2017. Besides purely gunfire based incidents, I note that this list includes attacks with the involvement of gun along with other weapons and violence (e.g., car ramming, butcher knife etc.).

Finding Campus-Specific Social Media Data Souce

As explained in the previous section, the social media data source of our study comes from online college communities (subreddits) on Reddit. Among the incidents involving gun violence on college campuses, I look for colleges which have subreddits with at least 500 subscribers on the day of incident on campus — the choice of this threshold is inspired form prior research that provides a rough estimate of number of unique subscribers in college subreddits that sufficiently represents the campus student population [31]. I use Reddit's subreddit search functionality feature, and retrieve number of subscribers from

[2]These sources include gunviolencearchive.org, time.com, motherjones.com, huffingtonpost.com, en. wikipedia.org; All Accessed: 2017-04-09

Reddit Metrics[3]. This leads to 12 such colleges meeting the criteria, and the number of subscribers in these subreddits ranges between 969 (r/NAU) to 8,936 (r/OSU).

Compiling Treatment and Control Data from Social Media

This study aims to to ensure the measured differences in stress after the gun violence incident is attributable to the incident, instead of other confounds and latent factors. In the statistics literature, the concerns around quantification of an "outcome" (stress) are typically mitigated by adopting randomized experiments, where, given a "treatment" (gun violence incident) in the target population, an equivalent population is assigned to a "control" (gun violence free) condition to rule out the effects on the outcome that are attributable to confounding or omitted variables [307, 477]. Given that an experimental approach would be infeasible and inappropriate in our case, I adopt a statistical matching technique, drawing from the causal inference literature [515]. To compile the dataset, for each of the 12 violent incidents, I identify two separate time periods of campus subreddit data collection:

***Treatment* Period.** I identify a two-month period before and two-month period after the gun violence incident on each campus. The rationale behind the choice of the duration of period of analysis stems from prior work [358], wherein it has been observed that effects of a societal upheaval persists a limited period of time. Because we focus on college campuses that tend to follow a 4 month semesterly or a 2.5 month quarterly academic system, I deduce that a four month period around each incident that closely follows the academic system would be able to glean meaningful stress changes that are attributable to the incident. This is labeled as the *Treatment* period.

***Control* Period.** For the combined period of two months before and two months after the gun-related violence incident on each campus, I identify an equivalent period of four months from the previous year. Gathering data from exactly the same period in the past year (when

[3]redditmetrics.com Accessed: 2017-04-09

Table 3.3: List of gun-related violence in U.S. college campuses during 2012-16 used in our work, along with the date, number of casualties (#n), and descriptive statistics of the corresponding subreddit communities.

College	Incident	#n	Subreddit	Users	#Posts
University of Southern California	2012-10-31	4	r/USC	1,143	2,676
University of Maryland	2013-02-12	3	r/UMD	2,201	9,578
University of Central Florida	2013-03-18	1	r/ucf	2,886	13,708
Massachusetts Institute of Technology	2013-04-18	3	r/mit	1,568	1,682
Purdue University	2014-01-21	1	r/Purdue	3,605	11172
University of California Santa Barbara	2014-05-23	21	r/UCSantaBarbara	3,278	17,682
Florida State University	2014-11-20	4	r/fsu	3,859	8,150
University of South Carolina	2015-02-05	2	r/Gamecocks	1,903	1,661
University of North Carolina at Chapel Hill	2015-02-10	3	r/chapelhill	2,025	1,177
North Arizona University	2015-10-09	4	r/NAU	969	1,025
University of California, Los Angeles	2016-06-01	2	r/ucla	6,301	9,454
Ohio State University	2016-11-28	14	r/OSU	8,936	35,372

no gun violence was reported) likely rules out confounding effects in the measuring temporal or linguistic differences in stress attributable to academic calendar factors, or seasonal and periodic events that impact students' experiences, lifestyle, and activities. Data collected from this period would minimize the confounds attributable to campus characteristics and student populations. This is labeled as the *Control* period.

I use Google BigQuery [256] to collect data from each college subreddit in *Treatment* and *Control* periods. This dataset contains 113,337 posts[4] (see Table 3.3). Further, each of *Treatment* and *Control* datasets are broken into *Before* and *After* samples based on whether the dates of posts are prior to or following the date of the reported gun violence incident (or the same date in the previous year).

Building a Stress Classifier

The first research aim is to measure stress manifested in the social media posts of different college campuses (RQ1). In the absence of ground truth labels on this data, I adopt a transfer learning approach, wherein I build a supervised machine learning model to classify stress expressions in posts into binary labels of *High Stress* and *Low Stress*. I use this classifier to machine label posts in the college subreddits.

[4]This study refers to 'posts' within college subreddits as a unified term for both posts and comments.

Class Definitions. This study's stress class definition is based on the established psychometric measure of stress as per the Perceived Stress Scale (PSS) [130]. The widely used 10-item version of PSS identifies three categories, scores ranging: 1) 0-13 are considered minimal stress, 2) 14-26 are considered moderate stress, and 3) 27-40 are considered extreme stress. However, typically, very few score in the extreme stress category – except for those who suffer from chronic stress challenges. Scores around 13 in the scale are considered average that typically split respondents' scores into two classes. Moreover, factor analysis [298] is known to reveal two factors, based on this scoring. This motivates our choice of two classes – *Low Stress* and *High Stress*.

Transfer Learning Data. I obtain all 1,402 posts from the *r/stress* from December 2010 to January 2017. The *r/stress* community allows individuals to self-report and disclose their stressful experiences and is a support community. For example, two (paraphrased) post excerpts say: "*Feel like I am burning out (again...) Help: what do I do?*"; and "*How do I calm down when I get triggered?*". The community is heavily moderated; so I consider these 1,402 posts as ground-truth data for *High Stress* posts. Next, I use a second dataset of over 100,000 random posts obtained by crawling the landing page of Reddit; this dataset was used in prior work as non-mental health related posts in developing mental wellbeing index for college campuses [31]. I employ this dataset as a source of ground-truth data for *Low Stress* posts to build the stress classifier. To approximately balance the two classes, I use a randomly sampled 2,000 posts from this dataset.

Establishing Linguistic Equivalence. In the transfer learning framework, I employ a training dataset obtained from non-college campus subreddits. I situate the rationale behind its appropriateness to build the stress classifier. The primary demographic of Reddit constitutes young adults [262], which is also the predominant demographic of college students. Therefore, we could anticipate linguistic similarities in the content of the training dataset and college campus dataset. To quantify linguistic equivalence between the two sources, I borrow methods from the domain adaptation literature [157] to conduct pairwise comparison of word vectors [34, 529]. This technique involves first constructing word vectors using frequently occurring n-grams in each source of data, and then calculating a distance metric, e.g., cosine similarity, to assess their linguistic similarity. Cosine similarity of word vectors is an effective measure of quantifying the linguistic similarity between two datasets [476], and a high value would indicate that the posts in the two datasets are linguistically equivalent.

To apply this method, I extract the most frequent 500 n-grams from our training dataset, and the same from the posts of every campus subreddit. Next, using the word-vectors of these top n-grams (obtained from the Google News dataset of about 100 billion words [417]), I compute the cosine similarity of the two datasets in a 300-dimensional vector space. I observe that the similarity ranged between 0.94 and 0.96, with a mean value of 0.95, providing significant confidence in our ability to use of the training data in building a stress classifier for college campus posts.

Classification Approach. On the above training dataset, I obtain features for the stress classifier. I use Stanford CoreNLP's sentiment analysis model to retrieve the sentiment of the posts. I obtain the top n-grams ($n = 3$) from the posts to be used as additional features. Then, using the sentiment label and the n-gram features, I develop a binary Support Vector Machine (SVM) classifier (with a linear kernel) for detecting *High Stress* and *Low Stress* in posts. I use this classifier to machine label all of the *Before* and *After* post samples shared in the *Control* and *Treatment* datasets associated with the 12 college campuses.

Quantifying Temporal Dynamics of Stress

Corresponding to RQ2, I propose a suite of computational techniques to assess the temporal changes of stress following gun-related violence on college campuses. Drawing from the time series analysis literature, on the above classified *High Stress* posts, I conduct analyses on : 1) *time domain* and 2) *frequency domain*.

Time Domain Analytic Approach

Normalized Temporal Variability of Stress. First, I examine the temporal variability of *High Stress* expressed in subreddit posts around each campus gun violence (*Treatment* data), and a similar period in the previous year (*Control* data). I aggregate posts shared per day, and then normalize the number of posts labeled as *High Stress* on each day. In order to assign weightage to the number of *High Stress* posts on a day as well as its proportion in this normalization, I employ a variant of the TF-IDF (term frequency-inverse document frequency) estimation technique: we multiply the proportion of *High Stress* posts in a day with squared root of their count on the same day. I obtain temporal variability in stress for both *Control* and *Treatment* datasets, spanning the *Before* and *After* periods. To reduce irregularities in time series, I smoothen the measures of the temporal variability of stress, using a non-parametric curve fitting regression method of $lowess$.

Then, I use 0-lag cross-correlation to assess how the manifestations of *High Stress* around incidents differ from the same in a comparable timeframe in the past. Cross-correlation estimates normalized cross-covariance function between two time series, and it is a mechanism to assess the relationship between a pair of time series signals [69].

Before-After Change Analysis. To quantify the degree of change in stress expressions around gun violence incidents on college campuses, I estimate changes in the trend of *High Stress* preceding and following the incidents. I compute z-scores of *High Stress* posts on each day for the *Before* and *After* samples in the *Treatment* dataset. z-scores quantify the standardized variation around mean value of a distribution and help estimating the relative changes in a time series data. Since z-scores do not rely on absolute values in a time series, it suits the analyses spanning across different periods of time (like in our case) when social media activities might vary. I compute an average change in z-scores between the *Before* and *After* samples by taking difference of the mean z-scores in the two samples. Finally, I fit linear regression models in *Before* and *After* samples to obtain the trend manifested by *High Stress* over time.

Frequency Domain Analytic Approach

Crises events can trigger disruptions in lifestyle, activities, and psychological dynamics of affected populations [397]. To understand how gun violence on college campuses disrupts the general periodicity of expression of *High Stress* in social media posts, I transform the time series of *High Stress* posts in the frequency domain. Frequency domain analyses are particularly suitable to study the rate at which the signal in a time series varies and therefore helps us assess its periodicity [646]. Using Fast-Fourier Transform (FFT) [79], I obtain the frequency distribution (measured in terms of days) of *High Stress* posts in *Before* and *After* samples in the *Treatment* dataset. I quantify the disruption in periodicities of *High Stress* posts by employing: 1) *spectral density analysis*; and 2) *wavelet analysis*. For the former, I compare the spectral density of two waveforms computed using periodograms [661], and for the latter, I compute symmetric mean absolute percentage (SMAP) difference between the peaks at the signal waveforms of the two samples.

Quantifying Linguistic Dynamics of Stress

Per RQ3, I assess linguistic attributes that characterize *High Stress* posts following the campus gun violence incidents, I adopt two forms of language analysis: 1) *psycholinguistic characterization*; and 2) *incident-specific lexical analysis*.

Psycholinguistic Characterization

I characterize the psycholinguistics of *High Stress* posts in *Treatment* data. I employ the well-validated lexicon called Linguistic Inquiry and Word Count, or LIWC [475]. Borrowing from prior work [358] to compare the *High Stress Treatment* posts belonging to the *Before* and *After* samples, I use the following LIWC measures for understanding the expression of psychological attributes in social media: 1) *affective attributes* (anger, anxiety, negative and positive affect, sadness, swear), 2) *cognitive attributes* (causation, inhibition, cognitive mechanics, discrepancies, negation, tentativeness, certainty), 3) *perception* (feel, hear, insight, see, perception), 4) *interpersonal focus* (categories: first person singular and plural, second person, third person, indefinite pronoun), 5) *temporal references* (future tense, past tense, present tense), 6) *lexical density and awareness* (adverbs, verbs, article, exclusive, inclusive, preposition, quantifier, auxiliary verbs, relative, conjunction), 7) *biological concerns* (bio, body, death, health, sexual) 8) *personal concerns* (achievement, home, money, religion) and 9) *social concerns* (family, friends, humans, social, work).

Incident-Specific Lexical Analysis

I examine the lexical cues shared in *High Stress* posts in *Treatment* data for each college campus. I examine to what extent, incident-specific language directly surface in the *High Stress* expressions shared in the college subreddits. I analyze posts shared within 7 days before and after the day of incident, to focus on weekly patterns. This chosen time window of analysis is inspired from prior work [492] that demonstrates that major shifts in psychological states and emotional responses are manifested until 7 days after the date of a crisis incident. Other work in social media analytics [163, 252] further indicates that human affective and mental health patterns follow stable weekly patterns, with systematic waning and intensification through different days of the week.. I extract 50 top occurring n-grams ($n = 1, 2, 3$) shared in the 7 days after the incident, and compute Log Likelihood Ratio (LLR) with respect to their occurrences in posts 7 days before the incident. The LLR for an n-gram is determined by calculating the logarithm (base 2) of the ratio of its two probabilities, following add-1 smoothing. Thus, when an n-gram is comparably frequent in the two week-long periods, its LLR is close to 0; it is closer to 1, when the n-gram is more frequent in the posts after the incident, whereas, closer to -1, for the opposite.

Table 3.4: Performance metrics of stress classification per k-fold cross-validation (k=5)

Metric	Mean	Stdev.	Median	Max.
Accuracy	0.82	0.11	0.78	0.90
Precision	0.83	0.14	0.77	0.92
Recall	0.82	0.09	0.78	0.88
F1-score	0.82	0.11	0.79	0.89
ROC-AUC	0.90	0.08	0.78	0.95

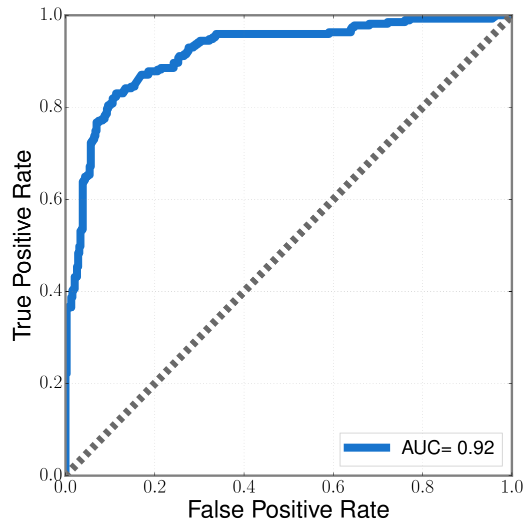


Figure 3.1: ROC curve of the stress classifier

3.2.2 Results

RQ 1: Inferring Stress in Social Media

Building a Stress Classifier

Corresponding to RQ1, I present the results of the machine learning classifier of stress. The binary SVM classifier uses 5000 n-gram features and three boolean sentiment features of Positive, Negative and Neutral; the number of n-gram features was determined based on systematic parameter sweep. I use a k-fold (k=5) cross-validation technique to evaluate our model, and achieve a mean accuracy of 0.82. This accuracy beats the baseline accuracy (based on a chance model) of 0.68 on this dataset. Table 3.4 reports the performance metrics of the stress classifier and Figure 3.1 shows the Receiver operating characteristic (ROC)

Table 3.5: Top 30 features in the stress classifier. Statistical significance reported after Bonferroni correction (*** $p < 0.001$).

Feature	p	log(score)	Feature	p	log(score)
stress	***	9.63	thank	***	6.20
try	***	7.46	meet	***	6.17
work	***	7.20	life	***	6.07
anxiety	***	7.05	sleep	***	6.03
meditation	***	6.88	problems	***	5.98
help	***	6.81	control	***	5.95
focus	***	6.62	job	***	5.89
luck	***	6.62	good	***	5.87
breathing	***	6.44	health	***	5.87
techniques	***	6.33	week	***	5.86
feel	***	6.30	minutes	***	5.83
exercise	***	6.30	doctor	***	5.83
time	***	6.25	mental	***	5.83
play	***	6.23	relax	***	5.72
body	***	6.21	stressful	***	5.67

curve of the same. The classifier yields low number of false positives (average precision 0.82), as well as low false negatives (average recall 0.82), indicating robust performance on test data.

Table 3.5 reports the top 30 features of our stress classifier. We observe a notable number of verbs or action-based nouns occur in this list, such as, *try*, *work*, *help*, *focus*. This also includes words contextually related to the expression of stress, like *stress*, *anxiety*, *stressful*, and *relax*. Aligning with prior work that has examined the correlates or factors precipitating stress [554], other notable words which occur in the top features include – 1) work-related: *work* and *job*; and 2) health-related: *health*, *body* and *sleep*.

Expert Validation of Stress Classifier

To understand the temporal and linguistic dynamics of *High Stress*, specific to our problem (RQ2 and RQ3), I apply the stress classifier to machine label the posts in both the *Treatment* and *Control* dataset. With the help of three human raters, expert in social media analytics and the study of affect dynamics, I validate a random sample of 151 of the classifier labeled posts (79 *High Stress* and 72 *Low Stress* posts). Our experts adopt the Perceived Stress Scale [130]

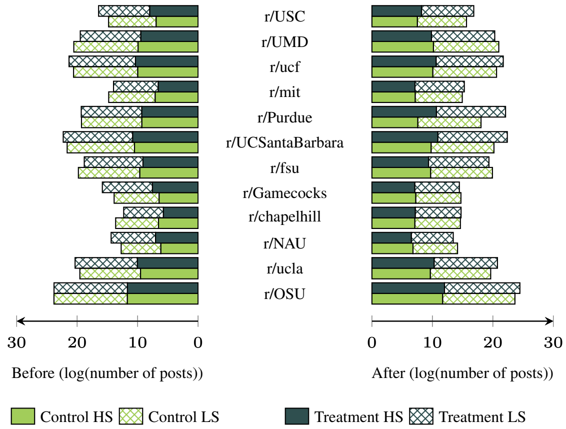


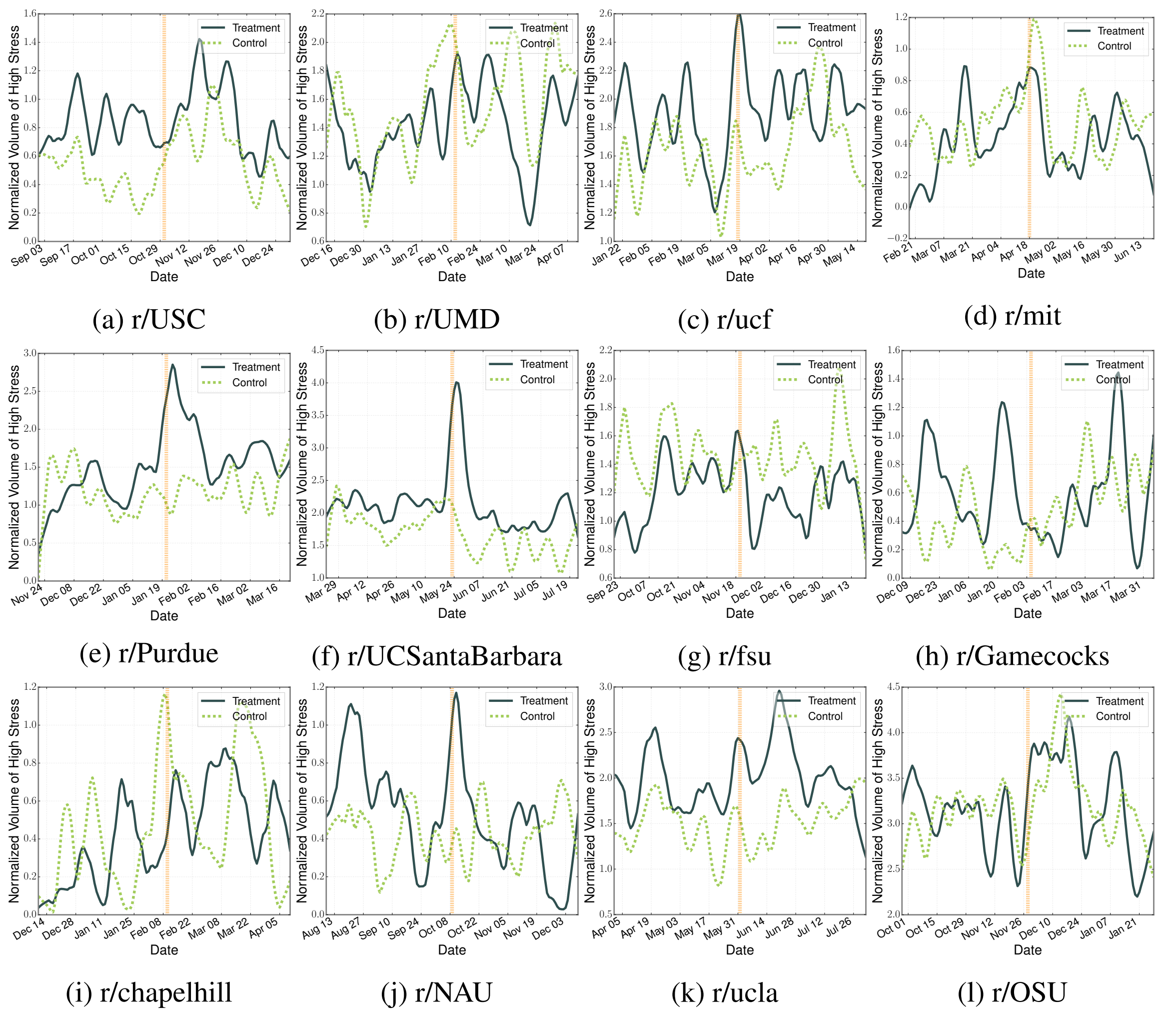
Figure 3.2: *High Stress* (HS) and *Low Stress* (LS) posts in *Before*, *After* samples of *Control*, *Treatment* datasets.

to examine how specific concerns measured in the scale (e.g., feelings of nervousness, anger, lack of control) are expressed in each post they rated. Presence of these concerns meant a *High Stress* label, while their absence indicated *Low Stress*. Our raters reach a high agreement in this task (Fleiss' $\kappa = 0.84$), and we obtain an accuracy of 82%[5] for the stress classification.

RQ 2: Temporal Dynamics of Stress

For RQ2, I summarize the results of class-wise stress distribution on each of the campuses in Figure Figure 3.2. I compare the *Treatment* and *Control* datasets spanning the *Before* and *After* periods to find: 1) for the *Treatment* dataset, the proportion of *High Stress* posts in the *Before* sample ranges between 35% and 45%, averaging at 40% (10,043 out of 24,737), whereas, the same for posts in the preceding year, ranges between 33% and 43%, averaging at 40% (9,415 out of 23,430); 2) the proportion of *High Stress* posts in the *After* sample of *Treatment* dataset, ranges between 33% and 47%, with a mean value of 41% (12,834 out of 31,370), while a similar period in the preceding year, reveals a mean proportion of 42%

[5]This should not be confused with cross-validation accuracy of stress classification. The same value in both the cases is only coincidental.



(a) r/USC (b) r/UMD (c) r/ucf (d) r/mit

(e) r/Purdue (f) r/UCSantaBarbara (g) r/fsu (h) r/Gamecocks

(i) r/chapelhill (j) r/NAU (k) r/ucla (l) r/OSU

Figure 3.3: Temporal variation in the expression of *High Stress*. The reference line represents the date of gun-violence incident.

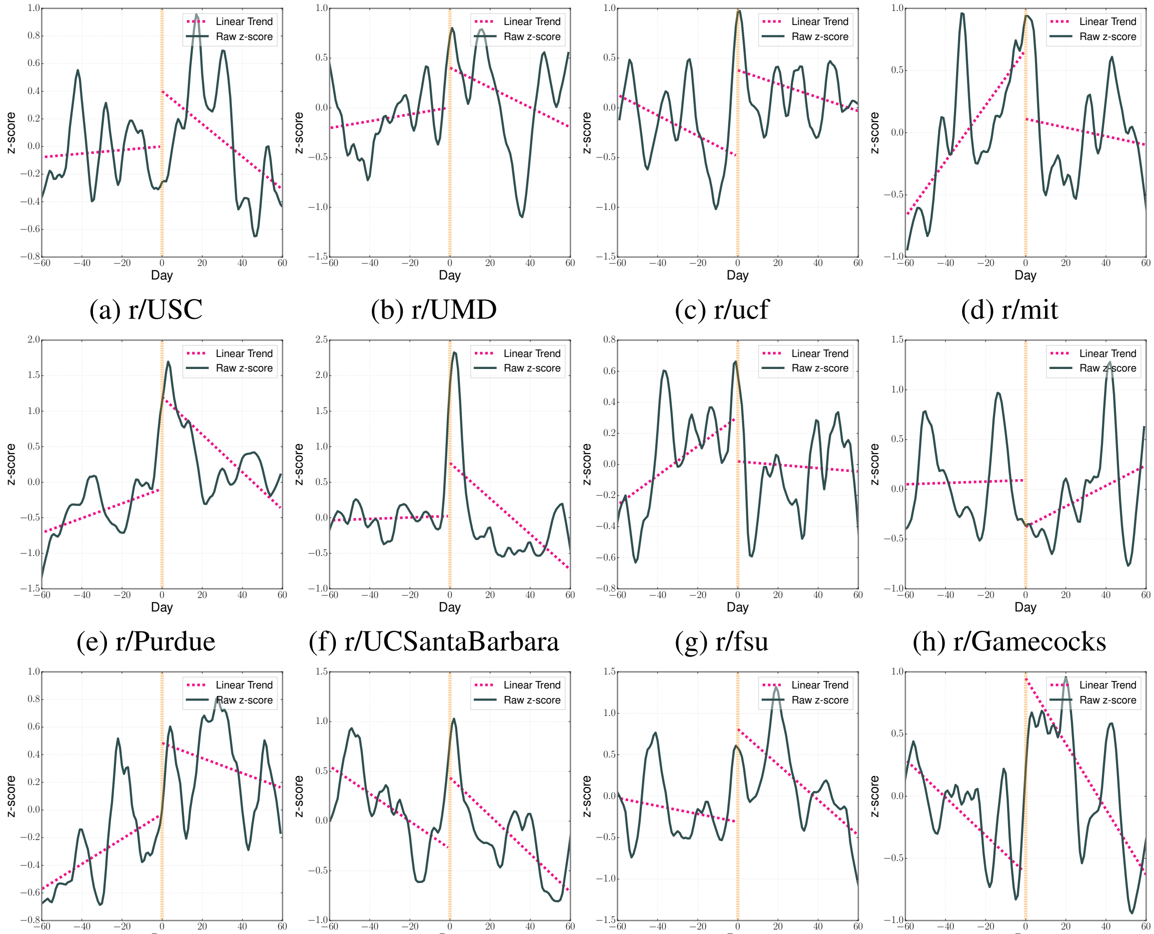
(9,528 out of 22,816). *These numbers convey that the proportion of posts expressing High Stress in the college subreddits remains comparably similar over an extended period of time, despite a gun violence incident on the campus.* However, as the ensuing time series analysis shows, we do observe significant changes in the *patterns of expression* of *High Stress* posts in the aftermath of gun violence.

Time Domain Analysis of High Stress Posts. To understand how stress varies following incidents of gun related violence on college campuses, I conduct time domain analysis of the expression of *High Stress* campus subreddits.

Temporal Variability of Stress. Figure 3.3 shows normalized volume of *High Stress* content in the *Treatment* and *Control* datasets. This study observes that *High Stress* posts are shared in the college subreddits all throughout the period spanning both the *Treatment* and *Control* datasets, in varying degrees. These posts consist of content ranging across varieties of academic and college-life specific topics including, admission, examination, or assignments: "*I really should be doing homework right now...*"; and "*I applied to the PhD program. I have emailed them twice in the past few weeks, but they keep saying they aren't done reviewing applications. [..] What should I do?*". This observation aligns with prior literature that situates various college-life specific factors to be attributable to student stress [518], and that stress is a persistent psychological observation among college students [50].

Specifically examining the day of the gun violence incident and its vicinity, this study finds a peak of the normalized volume of *High Stress* posts in a majority of the subreddits, considerably distinct in r/ucf, r/Purdue, r/UCSantaBarbara, r/NAU, r/OSU. The peak in stress in the *Treatment* year, as compared to the *Control* year supports a weak causal claim: that the campus gun violence contributes to an increased stress immediately following the incident. The mean normalized stress in the *Treatment* year is higher than the same for *Control* across all campuses (1.35 vs. 1.19), with the maximum difference observed in r/ucla (0.49) and r/UCSantaBarbara (0.45).

To assess whether the above reported differences are statistically significant, I conduct cross-correlation analysis of the temporal occurrences of *High Stress* posts in *Control* and *Treatment* datasets in Figure 3.6(a). We find negligible values of 0-lag cross correlations between the two time series, ranging between -0.002 and 0.003, with a mean value of 0.000 . This indicates that the differences between the pattern of *High Stress* between year of incident and a similar period in the year prior to it are indeed significant.



(a) r/USC (b) r/UMD (c) r/ucf (d) r/mit

(e) r/Purdue (f) r/UCSantaBarbara (g) r/fsu (h) r/Gamecocks

(i) r/chapelhill (j) r/NAU (k) r/ucla (l) r/OSU

Figure 3.4: Variation of z-scores of stress expressed in the *Before* and *After* samples in the *Treatment* dataset.

Before-After Change Analysis However, how does the expression of *High Stress* in the college subreddits change in the aftermath of the gun violence incidents, compared to that before? To answer this, I report the findings of our proposed before-after analysis.

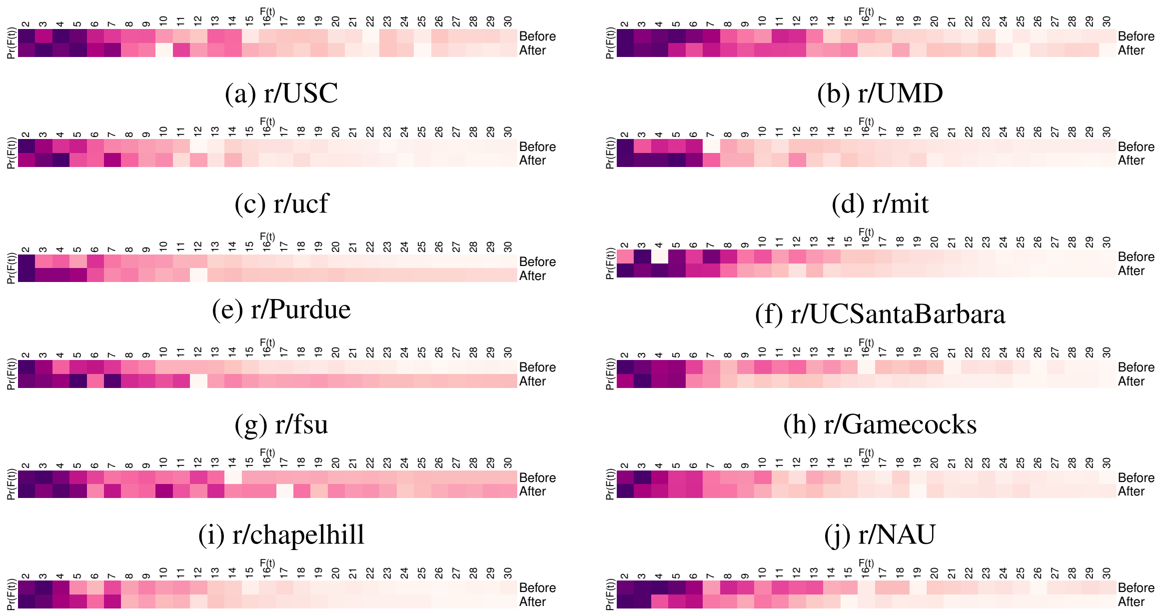
Within the *Treatment* dataset, I conduct a cross-correlation analysis between the temporal occurrences of *High Stress* posts, shared *Before* and *After* the date of incident. A 0-lag cross-correlation for *Before* and *After* samples (Figure 3.6(a)) ranges between -0.016 and 0.019, with a mean value of 0.005, reveals negligible correlation in the pattern of *High Stress* following the incident, as compared to before it. Next, I compute the z-scores of *High Stress* expressed on each day (Figure 3.4). At a glance, I observe the mean change in z-score between *Before* and *After* samples ranges from -0.30 (r/NAU) to 0.83 (r/Purdue), with 9 out of 12 subreddits exhibiting a positive change in expression of *High Stress* (Figure 3.6(b)). I conduct Mann-Whitney U tests of *Before* and *After* day-wise z-scores, for which I convert the dates to ordinals by counting the number of days on the either ends of the date of incident. These tests reveal statistical significance for each of the subreddits, with U ranging between 13,585 and 18,562,124, with a mean of 2,695,608.

More careful examination of Figure 3.4 indicates that the z-scores of *High Stress* in the days following the incident in most of the subreddits have a trend line (based on fitting a linear model) yielding a negative slope. Specifically, the most negative slopes are found in the cases of r/Purdue (-0.03) and r/OSU (-0.03). However, the trend line fits for *High Stress* z-scores in the *Before* period do not show such a trend– the mean slope during the period preceding the gun violence incidents is 0.001, revealing approximately a stable pattern.

Overall, these results suggest that the expression of *High Stress* in the aftermath of gun violence shows an abrupt shift in their temporal pattern, peaking significantly around the day of the incidents, and thereafter showing a downward trend.

Frequency Domain Analysis of High Stress Posts. The final analysis for RQ2 centers around understanding how the various gun violence incidents on campuses disrupt the periodicity of sharing *High Stress* posts. For this, working within the frequency domain, I apply Fast-Fourier Transform (FFT) on the distribution of *High Stress* posts in *Treatment* data. For each college subreddit, Figure 3.5 shows the distribution of frequencies $F(t)$ during *Before* and *After* periods as heatmaps. The color intensity of a cell in a specific heatmap indicates the probability of a certain frequency, $P(F(t))$ (measured in terms of days). Discussing the main observations from the heatmaps, in case of r/USC (Figure 3.5(a)), we find that *High Stress* posts in the *Before* period shows high periodicity (i.e., exhibit peaks in expression) around every 4 and 13 days, whereas the same in the *After* period occurred at every 5, 7 and 11 days.

To quantify if and to what extent periodicities of *High Stress* expression in the *Before*



(a) r/USC (b) r/UMD (c) r/ucf (d) r/mit (e) r/Purdue (f) r/UCSantaBarbara (g) r/fsu (h) r/Gamecocks (i) r/chapelhill (j) r/NAU (k) r/ucla (l) r/OSU

Figure 3.5: Frequency distribution heatmaps of stress in *Treatment* dataset. The x-axis, $F(t)$ represents frequency where t is in terms of days, and the density of color, $Pr(F(t))$, represents the probability of *High Stress* at $F(t)$.

and *After* samples, as given by the FFT approach, are disrupted around the gun violence incidents, I present the results of the spectral density and the wavelet analysis in Figure 3.7. For the former, the changes in mean spectral densities between the *Before* and *After* samples range from -2% (r/mit) to 96% (r/UCSantaBarbara), with a mean absolute difference of 37%. For the latter, I find that the symmetric mean absolute percentage (SMAP) differences between the frequency waveforms of the *Before* and *After* samples averages at 10, ranging between 2 (r/fsu) and 24 (r/Purdue). These results suggest that the periodicity of expression of *High Stress* was disrupted considerably following the incidents of gun violence on the 12 campuses.

Note that for r/Gamecocks, which shows aberrant pattern compared to other subreddits in the time domain analysis, according to its frequency domain analysis distribution heatmap (Figure 3.5(h)), there is a significant change in the periodicity of expression of high stress following the gun violence incident in the University of Southern Carolina (14% change in spectral density and an SMAP difference of 17).

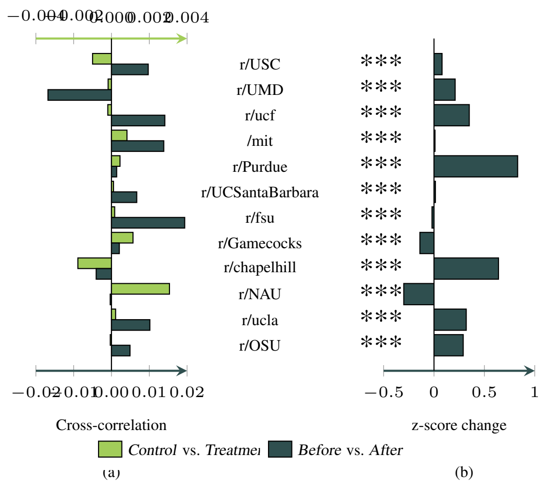


Figure 3.6: (a) 0-lag cross-correlation between *Control*-*Treatment* and *Before*-*After* (*Treatment*) samples. (b) z-score changes in *After* sample compared to *Before*. p-values computed using Mann-Whitney U-test (*** $p < 0.001$).

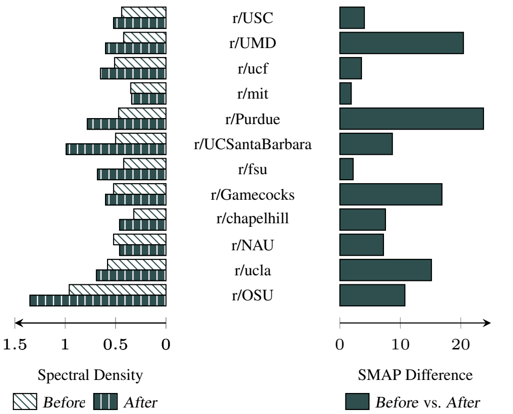


Figure 3.7: Spectral and wavelet analysis of frequency waveforms of *Before* and *After* samples in *Treatment* dataset.

RQ 3: Linguistic Dynamics of Stress

Finally, I present the results of our RQ3: the linguistic dynamics of *High Stress* expression in *Treatment* posts around the 12 campus gun violence incidents.

Psycholinguistic Characterization

In order to characterize the psycholinguistic cues of high stress content shared in the campus subreddits, I obtain the normalized occurrences of the LIWC attribute categories from the *High Stress Treatment* posts shared during the *Before* and *After* periods. For each psycholinguistic measure, to assess whether the differences between the *Before* and *After* samples are statistically significant, I perform Welch's t-test, followed by Benjamini-Hochberg-Yekutieli False Discovery Rate (FDR) correction. These results are presented in Table 3.6.

Table 3.6: Welch's t-test comparing the psycholinguistic attributes of *High Stress Treatment* posts shared *Before* and *After* gun violence incidents. Statistical significance reported after Benjamini-Hochberg-Yekutieli False Discovery Rate correction (*** $p < .001$, ** $.001 < p < .01$, * $.01 < p < .05$).

Category	*Before*	*After*	$\Delta\%$	t-stat.	p
Affective Attributes					
Anger	0.008	0.010	23.34	3.558	***
Anxiety	0.007	0.003	-61.81	-11.499	***
Negative Affect	0.007	0.009	20.55	3.376	***
Positive Affect	0.072	0.036	-50.56	-27.978	***
Sadness	0.002	0.002	14.42	1.554	*
Swear	0.006	0.007	12.46	1.5765	*
Cognitive Attributes					
Causation	0.027	0.013	-51.95	-23.312	***
Inhibition	0.008	0.005	-36.48	-7.824	***
Negation	0.029	0.041	41.90	13.334	***
Perception					
Feel	0.004	0.006	34.59	3.225	**
Hear	0.014	0.009	-35.49	-7.518	***
Insight	0.041	0.020	-50.24	-26.544	***
Percept	0.017	0.018	4.22	1.137	*
See	0.019	0.018	-7.11	-1.896	*
Interpersonal Focus					
1st P. Plural	0.013	0.010	-24.94	-5.011	***
1st P. Singular	0.061	0.080	32.47	15.864	***
3rd P.	0.015	0.012	-18.78	-3.740	**

Category	*Before*	*After*	$\Delta\%$	t-stat.	p
Temporal References					
Future Tense	0.037	0.035	-6.15	-2.15	*
Past Tense	0.056	0.061	8.58	3.79	***
Present Tense	0.116	0.113	-2.20	-1.79	*
Lexical Density and Awareness					
Article	0.117	0.144	22.93	16.72	***
Exclusive	0.032	0.064	99.31	33.66	***
Preposition	0.219	0.181	-17.38	-21.99	***
Quantifier	0.023	0.043	86.01	25.47	***
Biological Concerns					
Bio	0.012	0.014	10.62	2.48	**
Body	0.004	0.005	16.30	2.07	***
Health	0.003	0.007	97.39	8.84	***
Death	0.001	0.003	155.22	6.41	***
Personal and Social Concerns					
Achievement	0.037	0.016	-55.66	-25.51	***
Home	0.005	0.009	93.94	10.15	***
Money	0.022	0.011	-48.75	-16.66	***
Religion	0.003	0.004	43.82	2.87	**
Family	0.002	0.003	41.67	2.61	**
Friends	0.004	0.006	65.55	5.35	***

Affect Starting with *Affective Attributes*, we observe that *High Stress* posts in the *After* dataset show higher occurrences of *anger*, *negative affect* and *swear* words. Some example post snippets include, "*why the hell do they have a giant assault rifle?*" and "*I guess since campus is a gun free zone we're all fucked*". At the same time, *High Stress* posts in *After* period show significantly lowered levels of *positive affect* words. This indicates that the students may be engaging over Reddit to express their relatively higher negative perceptions, reactions and thoughts apropos the gun violence incidents.

Cognition and Perception For *Cognition* and *Perception*, we observe that, words related to *causation*, *inhibition* and *insight* are used significantly lesser in the *After* period. Prior work has related this psycholinguistic expression to lowered cognitive functioning [31] which is a symptom of high stress. However, *negation* words occur more frequently in the *After* period, and so do the words related to *feel*. Per prior work [475], this kind of greater perceptual expressiveness is known to be associated with language that depicts personal and first-hand accounts of real world happenings, events and experiences. Likewise, in this study, these changes indicate that the subreddit users are more expressive of their feelings in the aftermath of the campus gun violence incidents, e.g., "*I'm already home, but can't explain how am I feeling. No idea how to deal with this. Anything normal does feel way out of place at this point*."

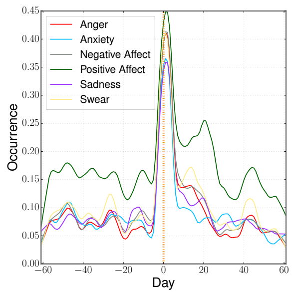
Linguistic Style Corresponding to linguistic style attributes, *Before* and *After High Stress* posts show distinctive *Interpersonal Focus*—we find that the use of *1st person singular* pronouns increases by 32% after gun violence, however that of *1st person plural* and *third person* pronoun words decreases. These patterns are known to indicate heightened self-attentional focus and greater detachment from the social realm [475]. Therefore, the users posting in the college subreddits may be resorting to social media to share their personal experiences and opinions about the incident. In the case of *Temporal References*, we find decreased use of *future* and *present* tense, and increased use of *past* tense in the *After* period. Higher use of past tense indicates tendency to recollect prior experiences and events [613], which in this case, might be an orientation to discuss the gun violence incident on the campus. With the exception of *preposition* words, all other function words (*exclusives*, *articles* and *quantifiers*) show significant increase in the *After* period, which are known to be related to a personal narrative writing style, often characteristic of crisis-inflicted populations [131].

Biological Concerns These results show that words referring all of *bio*, *body*, *health* and *death* increase in the *After* period. The *High Stress* posts shared following the gun violence incidents tend to relate to the after-effects, casualties, and implications of the incident for students' safety, well-being, and life (example post excerpt: "*I hope that no one is seriously injured or killed*").

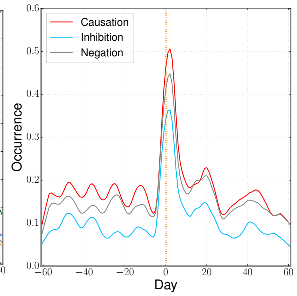
Personal and Social Concerns *Personal* and *Social* Concerns show revealing patterns. First, words related to *achievement* occur significantly lower (55%) in the period *After* the gun violence incidents. This category consists of words like 'confidence', 'pride', ‘progress’, ‘determined’. The college subreddits which are generally a platform for college-life discussions among students, lower usage of *achievement* words indicates a decline in tendency of engagements in career and academic topic related discussions in the aftermath of the campus gun related violence. Next, the usage of words relating to social life and relationships (such as, *family*, *friends* and *home*) increase after the incidents. This could indicate that the college subreddit users shared social wellbeing impressions and perceptions of solidarity in the context of the incidents. Words related to *money*, which occur significantly less frequently in the *After* samples, show that although money is a generic contributor of stress [612], it does not remain so in the *High Stress* posts following gun violence on college campuses. In addition, some of these incidents, such as the UNC Chapel Hill or the OSU attack were violence attributed to religious radicalism or religious hate-crime, which plausibly contribute to the higher occurrence of *religion* words in the *After* period.

Temporal Trends of Psycholinguistic Measures Finally, I examine the temporal trends of occurrences of the significant psycholinguistic measures in *High Stress* posts: ref. Figure 3.8. Most of the measures attain a peak in their occurrences in the immediate vicinity of the gun violence incident (day: 0). Interestingly, this peak is also the maximum value attained by all of the measures, with an exception of *feel* (Figure 3.8c).

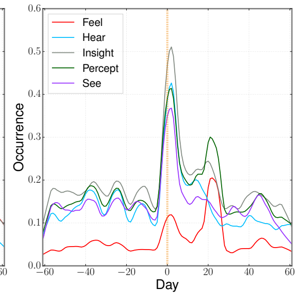
Comparing among the measures belonging to *Affective Attributes*, we notice that *positive affect* is expressed consistently higher than any other measure in this group. There is a revealing shift in trends of the usage of *Interpersonal Focus* (Figure 3.8d) — 1) The observations suggest a sharp increase in the usage of 1st personal plural pronouns, overtaking the usage of 1st person singular ones just immediately following the day of incident, which aligns with the emergence of collective identity as observed in prior work [397]. 2) But with days to follow, the occurrence of 1st person singular takes over, indicating a rise in usage of words referring to self attention. In addition, it is interesting to note the occurrence of *death*, agrees with prior work [241] — where although “death” has a minimal occurrence



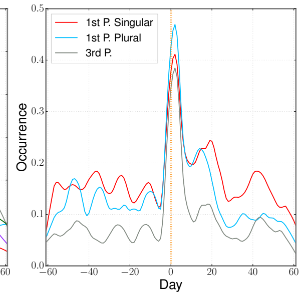
(a) Affective Attributes



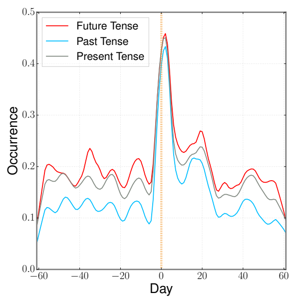
(b) Cognitive Attributes



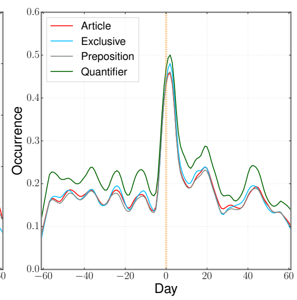
(c) Perception



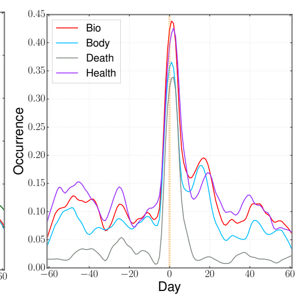
(d) Interpersonal Focus



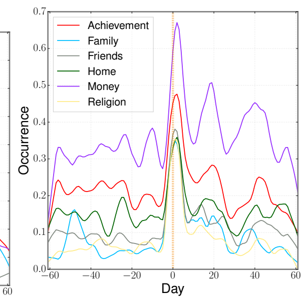
(e) Temporal References



(f) Lexical Density



(g) Biological Concerns



(h) Pers. & Social Concerns

Figure 3.8: Temporal variation of statistically significant psycholinguistic attributes of *High Stress Treatment* posts shared *Before* and *After* gun related violence.

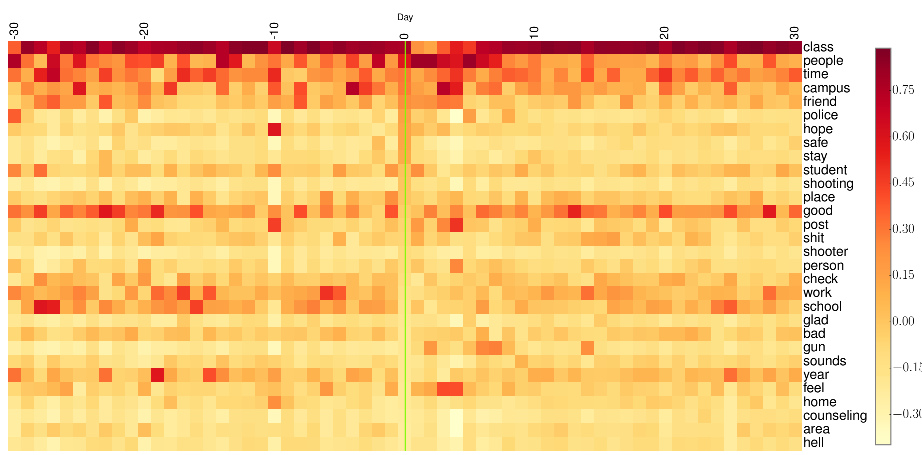


Figure 3.9: Top 30 keywords used in *High Stress* posts on the day of gun violence incident ($day = 0$) across all the subreddits.

consistently in the *Before* period, it achieves substantial concentration in *High Stress* posts just following the day of incident for a few days.

Table 3.7: Lexicon of selected n-grams ($n = 1, 2, 3$) occurring considerably higher in posts shared 7 days after the day of gun related violence, as compared to 7 days before.

Subreddit	***After*** $>$ ***Before*** **(**$LLR \geq (0.75)$**)**
r/USC	problems, night, security, shooting, party, events, fingerprint, entrances, email, dps, campus center, event, trojan, defense, safe
r/UMD	athletics, gun, supercar, cars, shoot, department, school, fire, community, sports, college, park
r/ucf	assault, assault rifle, weapon, tower 1, rifle, gun, police
r/mit	state, lincoln, stay safe, watertown, officer, officers, police, scanner, second, shots, shots fired, house, bpd, unknown, clear, confirmed, custody, dexter, fired, fuck, spruce, suspect, black, boston
r/Purdue	shooter, police, shooting, news, place, building, ee, campus, guy, heard, day, gt, know, student, people, today
r/UCSanta-Barbara	videos, victims, gun, mental, isla vista, guy, news, community, post, police, person, help, feel, love, iv, life, point, friends
r/fsu	mental, safe, shooting, strozier, ok, news, library, shooter, friends, victims, hope, stay, post, time, people, information, good
r/Game-cocks	alert, murdersuicide, public health, public health research, research center, shooter, shooting, students, support, lockdown, faculty staff, counseling center, building, health research center, cancelled
r/chapelhill	pretty, muslims, writing, religion, high, hicks, help, pound, students, execution style, execution, universal, world, abusalha, 30 serv, support, parking, unc
r/NAU	astronomy, jones, kill, kill people, meth, problem, harder, professors, self, self defense, shooter, shot, tour, guns kill people, year, guns, fight, class, defense, gun, asu, shooting
r/ucla	safe, confirmed, police, klug, shooter, gun, guns, health, mental, saying, professor, situation
r/OSU	safe, police, muslims, gun, removed, parking, post, stay, wrong

Incident-Specific Lexical Analysis

The final set of results includes analyzing linguistic markers as manifested in the subreddits immediately after gun violence on a college campus. For this, within the *High Stress Treatment* posts, I first extract the 30 most occurring n-grams ($n = 1, 2, 3$) on the day of incident. The 30-day *Before* and *After* temporal trends of usage of n-grams is shown in Figure 3.9 in a heatmap format. We note some contrasting patterns— for instance, 'class' occurs consistently in *High Stress* posts until the incident dates, but its usage declines considerably in the week following the gun related violence. On the other hand, subreddit users converse about 'people', 'friend', 'hope' and 'feel' a lot more in the immediate aftermath of the event, as compared to their overall occurrences—aligning with the observations drawn from the psycholinguistic analysis above, involving the emergence of a social orientation and greater perceptual expression following the incidents. The n-grams describing the nature, manifestation, and implications of the specific campus incidents, e.g., 'police', 'shooting', 'safe' and 'gun', have dense and increased concentration of usage following the day of the incident.

Next, I drill down further into the 12 colleges and, employ the Log Likelihood Ratio (LLR) measure to extract a lexicon of the top 50 n-grams ($n = 3$) from the *High Stress* posts within 7 days following the day of gun violence incident, and then compare their occurrence in *High Stress* posts in the 7 days preceding the day of incident. Table 3.7 reports the n-grams, for which n-grams with LLR of over 0.75, occur predominantly in posts after the incident, and are distinctive characteristic of what is discussed in the subreddits specifically in the *After* period.

Taking a close look at Table 3.7, we observe that this lexicon includes words which embed information specific to the incidents that occurred on the different college campuses under our consideration. For instance, we notice the lexicon to encompass words related to the geographical site of the incident such as – '*campus center*' in r/USC, '*tower 1*' in r/ucf, '*isla vista*' in r/UCSantaBarbara, '*library*' in r/fsu, '*public health*' in r/Gamecocks, and '*parking* in r/OSU. Next, we note the presence of the word '*videos*' in r/UCSantaBarbara and '*murdersuicide*' in r/NAU, which are coherent with how the incidents unfolded at these campuses. Additionally, agreeing with the findings from the psycholinguistic characterization presented above, we observe the presence of '*muslims*' in r/chapelhill and r/OSU, where the incidents were attributed to be religious hate-crimes or radicalism. We find the usage of words relating to the victim or the perpetrator's name and occupation in some of the subreddits, such as r/mit, r/Gamecocks, r/chapelhill and r/ucla. Summarily, this analysis shows that high stress expressed in posts of the college subreddits in the immediate aftermath of the gun related violence may be a consequence of the incidents in the respective campuses.

3.2.3 Discussion

I employ a causal inference based analytical approach, in conjunction with computational techniques to examine the evolution stress following gun violence events on college campuses. The findings suggest that, compared to a control (gun violence free) time period on each campus, there was a significant change in the volume of posts expressing high stress following the violent incidents, including a considerable change in the patterns of stress expressed in the immediate aftermath of the incidents. Then, psycholinguistic characterization of the high stress posts indicates that campus populations exhibit reduced cognitive processing and greater self attention and social orientation, and that they participate more in death-related conversations. Additionally, a lexical analysis of high stress posts shows distinctive temporal trends in the use of incident-specific words on each campus, providing further evidence of the impact of the incidents on the stress responses of campus populations.

I situate the findings in the context of psychological theories surrounding trauma and crisis. I derive two major observations: 1) Psychological stress may be automatically inferred from social media content by employing supervised learning approaches; and 2) Inferred stress levels in a college campus may indicate the responses of individuals exposed to the reported gun-related violence incident. To arrive at these findings, this study make a methodological contribution in this study as to how stress changes, temporal and linguistic, can be measured following a violent incident on campus, drawing from machine learning and time series analysis techniques. Therefore, this study bears implications for researchers intending to study the socio-psychological responses of a population exposed to a crisis, and those interested in developing technologies to assist vulnerable populations following traumatic events. I discuss these implications below.

Theoretical and Psychological Implications. Freud's psychoanalytic theory [228] argued that external reality, for example, traumatic events, can have profound effects on an individual's pysche, and can be considered to be the cause of emotional upheaval, stress, and traumatic neurosis. He suggested that the personal impact of the trauma, the inability to find conscious expressions for it, and the unpreparedness of the individual can cause a breach to the stimulus barrier and overwhelm the defense mechanisms [227]. This study examines these theoretical constructs in a data-driven manner. For instance, the linguistic analytical methods suggest distinctive psycholinguistic cues in high stress posts shared after the gun violence incident compared to before. As an example, language related to biological concerns increase remarkably following the incident. In contrast, more general topics closely related with stress in a college population, such as financial and career-related concerns [518] exhibit significant reductions following the incidents.

Further, the incident-specific lexical analysis makes a notable finding — the content shared on social media immediately following the violent incidents are largely topically related to the events. McCann and Pearlman [407] adopted the framework of cognitive theory to propose seven fundamental psychological need areas following experience of a crisis event: frame of reference, safety, dependency/trust of self and others, power, esteem, intimacy, and independence. Trauma, they argue, may cause disruptions in any of these need areas and thereby lead to troublesome emotions and thoughts such as stress. Words such as "stay safe", "support", "hope", "help", "self", that increase in usage in high stress social media posts following the incidents, indicate the expression of many of these needs.

These methods also help uncover the nuances in acute stress responses on college campuses following the violent incidents, that tend to offset more persistent chronic stress expressions. For instance, although students undergo stress all throughout the year because of academic and personal reasons [518], stress expression of a campus changes considerably after gun violence. In essence, our findings show that, as revealed by campus social media posts, stress as a construct, is prevalent (possibly chronic in nature) across time, yet the nature of this construct changes drastically (possibly turning more acute) around a critical crisis incident. This also reveals temporal and linguistic "signatures" of expression of such acute stress, such as altered periodicities or increase/descrease in specific psycholinguistic words can be gleaned with our proposed machine learning and time series analysis approaches. These findings support similar observations made with respect to the manifestation of psychological states in a response to chronic violence [168], wars [397] and terrorist attacks [377]. Moreover, closely aligning with prior work [23], we also observe that the post violence acute stress levels subside within days to follow, and approach baseline levels of generally persistent chronic stress. This could be because life goes back in order, with other aspects of campus life taking priority. This interpretation is consistent with prior work in crisis informatics [397], as well as Foa and colleagues' emotional processing theory [221]. They noted that emotional experiences, such as anxiety and stress, are often relived well after the original traumatic events have occurred, although the frequency and the intensity of emotional reliving usually decreases over time.

Practical Implications. The computational techniques provide a robust mechanism to quantify the impacts and severity of a crisis, as well as the corresponding community responses. These techniques will find use as unobtrusive sensors of stress and its linguistic and temporal changes during crises. These methodologies may be leveraged in future situations where causes of stress may not be so apparent or known, as was the case in our study, e.g., assessing stress and associated student responses in everyday (crisis free) contexts, where a variety of day-to-day but unanticipated academic, personal or social concerns may contribute to stress.

The social media of specific affected communities (college campuses) can help identify unique "signatures" or idiosyncratic patterns in stress expressions. The temporal trajectory of high stress following the incident at the University of Southern California was distinct from the others; so were the kinds of linguistic cues that surfaced in social media content immediately after the incidents. These approaches may help discover the presence of protective factors surrounding stress in specific communities, including how a campus's stress expression deviates from the expected pattern of stress on any campus affected by a similar crisis. This information can be valuable to crisis rehabilitation efforts, including how specific campuses may adopt policies or strategies to enhance the idiosyncratic aspects relating to the community, that exacerbate or protect against stress.

The impacts of a violent incident transcend observed casualties, and could be perceived very subjectively at individual level. This work provides a way to account for the "invisible wounds" [306] or "hidden casualties" [491] in a crisis, which tend to not be measured adequately. In essence, this study observes that in the aftermath of campus gun violence, campus-specific social media like Reddit, acts as a unique platform allowing campus populations to express emotion and stress about the circumstances, (semi) anonymously, amid feelings or perceptions of fear and trauma. These techniques enable capturing a "quantitative narrative" of the self-disclosed stress experiences of campus populations exposed to crisis events, which can eventually inform historical accounts about campus life.

Design Implications. This research bears implications for designing technology that can support improving mental health provisions for campus populations during times of upheavals. A significant challenge for college administrators is providing adequate mental health services, such as around aggravated stress levels, in a proactive, real-time manner. These efforts become even more difficult in the face of campus crises, due to the disruption in everyday life and activities on campus. This work shows promise in enabling technology-assisted means to tackle these challenges.

Sudden bursts of stress can be detrimental in the long-term [409]. Our work can help build population-centric stress tracking tools. These tools can advance the current practices in how college authorities engage with the student community following crisis incidents. Typically these practices include broad campus-wide communication of the context and outcomes of the incident, followed by specialized programs to direct psychological counseling and rehabilitation support to students who may need help. This work can complement existing techniques and tools for assessing stress among individuals [688]. These tools can help inform the college authorities to learn about the pervasiveness of stress following crises and the degree of disruption from normalcy. This can enable the college stakeholders in making empirically informed decisions about the nature of crisis communication that should take place on campus, such as balancing informational alerts with adequately sensitive and focused assurance. Additionally, administrators can gather a better understanding of students' counseling or rehabilitation resource needs. They can identify specific stress induced temporal or linguistic responses that negatively impact specific student groups. This can allow them to take adequate action in a timely manner, e.g., conducting campus-wide awareness and mitigation campaigns on mental wellbeing, or making tailored provisions to improve mental resilience and morale of the student body.

Next, often following campus violence, administrators need to make policy decisions for maintaining the safety and morale of the student body. For example, in the wake of 2012 shooting in University of Southern California (which was considered in our analysis), the administration incorporated heightened security measures and visitor restrictions in the campus[6]. The patterns and observations gleaned from tools that leverage our methods, such as the incident-specific linguistic markers, could provide evidence-based, student-contributed insights to administrators so as to make informed policy decisions to scaffold campus life succeeding crises. Over time, this ability to identify markers of student stress and their dynamics can also contribute to improved preparedness in campus around future crisis events.

3.3 Examining the Effects of Counseling Recommendations After Student Deaths on College Campuses

Crisis events on college campuses can have a profound negative impact on the overall wellbeing of the campus community [608]. One such crisis that is frequently encountered is the death of a student. Recent statistics report that two in every 1000 U.S. college students die every year, because of accidental, suicidal, and acute and chronic illness reasons [625]. Among these, campus suicides have almost tripled within the last fifty years, and about 18% of undergraduates and 5% of graduate students have had lifetime thoughts of attempting a suicide [604]. These alarming statistics not only hint at the strains of campus and academic life, every such tragic incident also has widespread repercussions by affecting

[6]latimesblogs.latimes.com Accessed: 2017-04-15

the general psychological wellbeing of the campus [669]. In fact, some of the most dangerous consequences of such crises include “copycat suicides” (when student suicides come in clusters due to social contagion effects) [483], or heightened risk to serious mental health challenges like post-traumatic stress disorder [549]. Given students already under-utilize mental health care resources due to social stigma, lack of awareness, and the pressures of academic life [196], unanticipated crises like student deaths bring additional challenges to the mental health amelioration efforts on campuses.

Crisis events on college campuses, such as student deaths, therefore, underscore the necessity to reinforce existing intervention programs or undertake new initiatives toward reducing the psychological effects of the crisis in the student community [66]. A common approach adopted by campus administrators involves public communication and outreach, promoting information about various student-centric support, coping resources, and counseling services. Given the pervasive use of web-based technologies in the college student demography [479], these recommendations are often shared via email and social media, also because such communication channels bear the potential to provide a common, stigma-free platform to comment and discuss about the event itself, as well as to grieve and cope. Figure 3.10 shows an excerpt of one such post shared by a campus administration on Reddit. This study refers to such posts as “*counseling recommendations*.”

However, significant methodological gaps exist in measuring the effectiveness of these post-crisis interventional recommendations shared by campus officials [563]. These range from a reliance on retrospective self-reports, to the difficulty in causally determining the link between exposure to these recommendations and the psychological states of students following a crisis [172].

I address the above gaps in examining the efficacy of counseling recommendations following student death incidents on college campuses, targeting two innovations. First, I use unobtrusively gathered social media data of college Reddit communities, where these recommendations are shared by campus officials. Social media helps us track individuals

Dean [redacted] letter to Students. self.gatech

Submitted [redacted] No_Flair_Selected

Students, staff, [redacted]ulty of Georgia Tech,

The sudden and tragic death of an undergraduate student, [redacted], has been devastating news to the Georgia Tech community, and in particular to all who knew [redacted]. For members of the [redacted] the [redacted], and for faculty who had him enrolled in their classes, the shock and grief are particularly acute.

A remembrance ceremony has been planned for [redacted]. in the lobby of the Ferst Center for the Arts. Counselors from the Georgia Tech Counseling Center and campus chaplains will be on hand at this event.

We are committed to providing resources for the mental, emotional, and physical well-being of our entire campus community. Please remember that Georgia Tech offers multiple services and resources in support of the community during this time of loss and grief:

· The Georgia Tech Counseling Center (http://www.counseling.gatech.edu) is staffed by psychologists and mental health counselors. They offer brief, confidential counseling and crisis intervention services to students. The Counseling Center also offers an after-hours on-call counselor to speak and consult with students in crisis. In addition, they sponsor a series of workshops for managing stress. · Stamps Health Services (http://health.gatech.edu/services/Pages/Psychiatry.aspx) is open to students and their spouses. Students interested in scheduling an appointment may call 404-894-2585 or visit the second floor of the Stamps Health Services Building. · Division of Student Life and the Office of the Vice President and Dean of Students has a referral option if you are concerned about a student (http://www.studentlife.gatech.edu). · The Georgia Crisis & Access Line (1-800-715-4225) is staffed with professional social workers and counselors 24 hours per day, every day, to assist those with urgent and emergency needs.www.mygcal.com. · Georgia Tech has contracted with EAP Consultants, LLC to provide employees and their family members with a comprehensive Employee Assistance Program (EAP). All healthcare benefits eligible employees and their families have access to the program. (http://health-and-wellbeing.gatech.edu/eap)

It is our hope that anyone who needs these services will be able to take full advantage of them. At times like these, we are reminded of the importance of coming together in support, understanding, and care for one another.

[redacted] Vice President and Dean of Students

Figure 3.10: An excerpt of a counseling recommendation post shared on *r/gatech* following the death of a Georgia Tech student.

who engage with these recommendations and what effects they have on their psychological states. Then, as a second innovation, I develop a causal analysis framework that statistically models the shifts in psychological states characterizing individuals who are exposed to these recommendations, and those in a control group. As indicators of these changes, drawing from natural language analysis (word embeddings) [417], the crisis literature, and psychological theories like the "grief work hypothesis" [562], I develop the following categories of measures: a) affective changes, specifically around the expression of grief (I model a new "grief lexicon"), b) behavioral changes, and c) cognitive changes.

Focusing on a dataset of ∼400M posts and ∼350K users spanning 174 college communities on Reddit, the findings show that, compared to baseline scenarios, in the aftermath of student death incidents, individuals who are actively exposed to the recommendation (via commentary) tend to show statistically significant shifts in their psychosocial attributes compared to a matched control cohort who do not engage with the recommendations in the same manner. Examining these changes, I find that the exposed group demonstrates greater expressivity of grief, shows signals of social integration and diversity in interactions, and exhibits improved cognitive processing as well as linguistic and stylistic complexity. This study situates the findings in the crisis and mental health literature that associate such shifts with a healing response, which in turn are indicative of benefits to one's psychological state. *This study provides the first large-scale, (social media) data-driven study of the effects of post-crisis counseling interventions.*

3.3.1 Data and Methods

Data

For this study, I use Reddit as our data source. Below I describe the approach to collect the datasets for our study.

Counseling Recommendations (CR) Dataset. Starting with a seed list of generic and campus-specific keywords, I first use an iterative snowballing technique to build a list of search queries to identify counseling recommendation posts in the 174 subreddits: *1) Generic Keywords* are related to death and counseling, such as *"death", "suicide", "counsel*", "rip", "therapy"*. This list also includes phrases related to email, and positions of responsibility, like *"email", "email dean", "president"*, *2) Campus-specific Keywords* are specific to a campus, which I compile by consulting the official college websites to obtain names of the campus administrators (e.g., president or dean) and the counseling body. Using these keywords, I query Reddit's search interface for counseling recommendation posts, and manually inspected the returned posts for correctness in terms of this study's definition of counseling recommendations. This leads to 88 counseling recommendation posts across 46 subreddits, which I denote as the CR dataset.

Table 3.8: Datasets on student death (SD) and counseling recommendation (C).

	SD	$\neg SD$
C	$\boldsymbol{CR}$	$\boldsymbol{B_2}$
$\neg C$	$\boldsymbol{B_3}$	$\boldsymbol{B_1}$

Baseline Datasets. Additionally, towards quantifying the psychosocial changes attributable to the counseling recommendations following student death events instead of other hidden factors (e.g., changes associated with active participation in *any* content shared by campus officials, exposure to content around non-crisis events, or general interest in counseling related content), I consider three other baseline datasets (ref. Table 3.8).

Baseline Dataset $\boldsymbol{B_1}$ includes announcements from campus officials unassociated with a crisis (student death) event and without any pointers to counseling or support resources. E.g., B_1 contains posts about non-crisis/non-critical campus events, and appointments or resignations of officials.

Baseline Dataset $\boldsymbol{B_2}$ consists of campus announcements unassociated with a student death but points to counseling services. E.g., it includes counseling recommendations that are either routine, or about socio-political issues and policies (e.g., immigration), sexual harassment, or violence on campus.

Baseline Dataset $\boldsymbol{B_3}$ includes posts that are campus announcements acknowledging a student death but *without* pointers to counseling information.

I acquire these datasets employing similar technique as in the case of CR posts—identifying keywords iteratively (e.g., *"sexual", "violence", "immigration", "policy", or "student affairs"*), querying and manually inspecting the correctness of returned posts. Eventually, B_1 had 229 posts, B_2 had 30 posts, and B_3 had 1 post across the 46 subreddits in which at least one CR post was present.

Next, I use nested queries on the cloud platform, Google BigQuery [256], which hosts an entire archive of Reddit data (Bagroy et al. 2017), to obtain the usernames of users who commented on the CR, B_1, B_2, and B_3 posts. I also collect these users' historical archives (or“timelines”) with all posts. Additionally, I collect similar data of 358,871 other users (378,381,052 timeline posts), who posted on the campus subreddits, outside of the CR, B_1, B_2, and B_3 posts. As a measure to restrict the corpus among those individuals who belong to the same campus per subreddit, I further prune the dataset of any users who posted on more than one campus subreddit. Finally, I identify 842 users and 3,167,266 timeline posts for the CR dataset, 2,215 users and 6,818,873 timeline posts for the B_1 dataset, 321 users and 1,231,784 timeline posts for B_2, and no users in B_3.

Measuring Effectiveness of Counseling Recommendations

Now, I present the measures via which I quantify the psychological effects of counseling recommendations. These measures are based on the three core psychosocial constructs elucidated in the psychology literature: a) Affective, b) Behavioral, and c) Cognitive attributes [77]. Inspired from the widely adopted “difference in difference” technique in the causal-inference research [1], I estimate the effects of counseling recommendations in terms of the changes corresponding to all the psychosocial measures in the *Treatment* and *Control* groups *Before* and *After* the date of a specific CR, B_1, or B_2 post.

Affective Changes

Research has demonstrated affective variability in individuals following crisis events [241, 377, 397]. This work models affect from the perspective of “grief”. Grief is a “response” and a mix of conflicting feelings and a wide range of strong emotions [325]. When someone dies, alongside bringing shock, disbelief, and numbness, it leaves friends and relatives feeling lost, anxious, depressed, or physically unwell [241]. Grief is the process by which I adjust to the death of someone close [549, 669]. A rich body of literature in psychology, by way of the “grief work hypothesis” [562] therefore has identified the coping and healing benefits of grieving [92], which in turn are associated with achieving timely resilience and return to normalcy and day-to-day activities following crises. Thus examining grief as a measure of

Table 3.9: Top 30 n-grams used discriminatingly in reddit grief communities. These n-grams are obtained by ranking their Log Likelihood Ratio (LLR) measures with generic non-mental health communities ($-1 \leq LLR < 0$), tf-idf values are scaled at 10^{-2}.

Word	tf-idf	Word	tf-idf	Word	tf-idf
thank	14.5	loved	4.65	help	2.82
sorry	13.0	husband	3.54	memories	2.72
loss	8.95	support	3.34	feelings	2.52
remove	6.77	passed	3.25	easier	2.51
hope	6.73	hugs	3.21	miss	2.37
lost	6.00	beautiful	3.13	son	2.36
grief	5.98	sharing	3.15	peace	2.34
death	5.84	glad	3.00	cancer	2.30
died	5.46	suicide	2.96	comfort	1.91
pain	4.94	heart	2.88	sucks	1.79

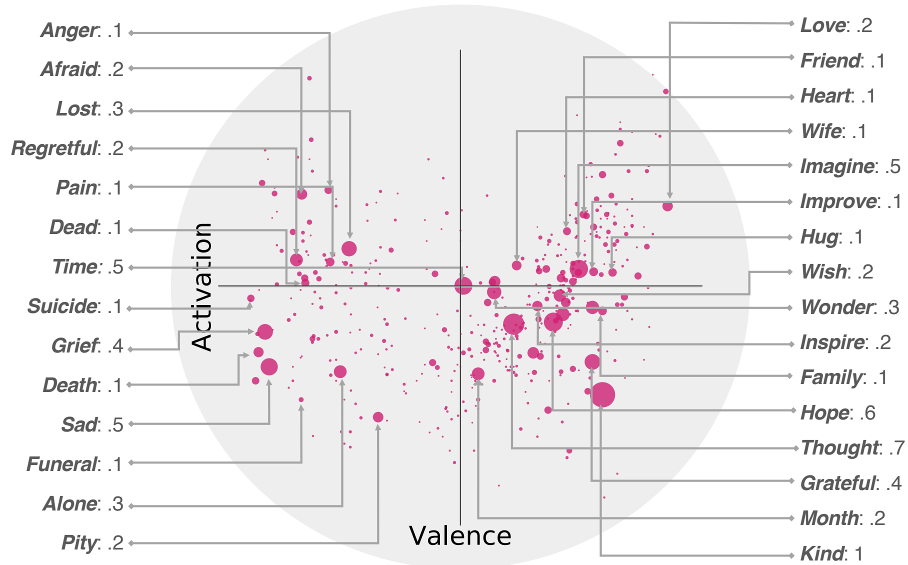


Figure 3.11: Weighted distribution of affect categories (ANEW) for grief lexicon on Russel's circumplex model. Top ANEW categories and their standardized tf-idf ($[0, 1]$) are labeled.

psychological change following CR exposure is extremely relevant in this study's setting.

While prior work has developed methods to identify affective attributes like mood, emotion, and sentiment [160, 529], presently, there are no computational means to infer grief from language. Moreover, due to the complexity of grief as an affective construct (note the definition above), gathering high quality ground truth is challenging. Furthermore, in assessing psychosocial changes among individuals particularly in response to an environmental stimulus (such as crisis), psychology literature and theories advocate a grounded representation of affect, comprising of not only the commonly used valence (pleasantness dimension), but also the intensity of affect, known as activation. To address these challenges, and to obtain a theoretically valid assessment of grief around the sharing of counseling recommendations, I employ a novel open vocabulary approach of 1) building a grief lexicon; and 2) mapping the words in the grief lexicon to two affective dimensions, valence and activation, drawing on the established Russell circumplex model of affect [487].

Building a Grief Lexicon. To build a grief lexicon, I adopt an open-vocabulary based transfer learning approach. I leverage data from 15 subreddits around the topic of grief, such as *r/grief*, *r/GriefSupport*, or *r/bereavement*, where people engage in sharing their sorrow and grieve about the loss of their loved ones. From these subreddits, I obtain over 50K posts (D_G), based on the archives available on Google's Big Query. Additionally, I obtain a generic Reddit corpus, D_R of posts unrelated to any grief or mental health issues, also used in prior work [31].

Thereafter, I extract all n-grams ($n = 2$) from the above two datasets D_G and D_R, along with their $tf\text{-}idf$ scores. Then, I use Log Likelihood Ratio (LLR) measures to obtain a ranked list of most distinguishing n-grams across the two corpuses. LLR for an n-gram is determined by calculating the logarithm (base 2) of the ratio of its two probabilities, following add-1 smoothing. Based on the LLR measures, when an n-gram is comparably frequent in both the datasets, its LLR is close to 0; it is < 0, when the n-gram is more frequent in D_G, and > 0 for the opposite. Among the 4,714 n-grams exhibiting negative LLR, I obtain a list of those 50% of n-grams with the most negative values— here, I use median as the measure of central tendency. These 2,357 n-grams with a big negative skew in LLR are most distinctive of D_G, and I refer to them as the "**Grief Lexicon**", L_G. Table 3.9 reports a sample of the top 30 of these n-grams ranked on their $tf\text{-}idf$ scores.

Modeling the Affective Dimensions of Grief. To characterize valence and activation dimensions of words in the grief lexicon based on the circumplex model, I employ the widely used *word embedding* technique to derive latent semantic relatedness between words [418] and the *Affective Norms for English Words (ANEW)* lexicon [440]. ANEW is an affect dictionary, curated after extensive and rigorous psychometric studies, containing a list of over 1,000 affect categories and their quantified measures of valence and activation. Prior research has successfully used ANEW to understand expression of mood and affect [160].

For every affect category in ANEW, I obtain its vector representation in a 300 dimensional word-embedding space using the word2vec model (pre-trained on Google News dataset of $\sim$100B words). Within the word-vector space, semantic similarity between any two words can be estimated with cosine similarity, using which I map all the n-grams in grief lexicon (L_G) to the most similar ANEW category (if any, threshold = 0.69 [502]) and obtain their valence and activation values. Accordingly, 2,357 n-grams from grief lexicon are mapped to 459 unique ANEW categories. With their valence and activation values as coordinates on an x-y frame and tf-idf as the magnitudes, I model the grief lexicon in the two-dimensional circumplex space of affect (see Table 3.11). I find that expressions across a range of valence and activation values occur frequently in grief, e.g., “kind”, “inspire”, “love”, “anger”, “sad”, “afraid”, and so on. This aligns with the definition of grief [325], and justifies this work’s lexically induced open-data strategy of modeling grief in the circumplex model of affect.

Characterizing *Treatment* & *Control* with Grief. With the above grief lexicon and its 2-dimensional affective model, I quantify the affective expression of grief in the *Treatment* and *Control* groups around the date of the CR, B_1, or B_2 posts in their respective datasets. Specifically, within each of these groups, I obtain all the n-grams and their tf-idf values before and after the date of post. Applying the same word-vector based similarity metric described above, I map these n-grams to the most similar grief word and its valence and activation value. I compute the mean percentage change of valence and activation of grief in the *Treatment* and *Control* groups.

Behavioral Changes

Next, I measure psychosocial changes in behavior around the date of counseling recommendation posts. In the psychology, mental health, and crisis literatures, many behaviors including changes in social functioning and shift of interests can be indicative of an individual's changing psychological trajectory. This study is interested in observing the following changes as effects of exposure to counseling recommendations: Does the user become more active on Reddit, indicating improved extroversion? Do they participate in more subreddits, indicating a diversity of interests and interactions? Do they involve themselves in more discussion threads on Reddit, indicating social engagement? Inspired from prior work [665], this study answers these questions with three metrics, a) activity, or frequency of posting, b) interaction diversity, that is, number of unique subreddits they participate in, and c) interactivity, given by computing the number of comments to post ratio.

Cognitive Changes

Literature in psychology identifies cognitive attributes as another indicator of an individual's psychological state [36, 388] —an uptick in wellbeing is known to be associated with reduced cognitive impairment and improved perceptual processing. Further, psycholinguistics literature has revealed the association of linguistic structural and stylistic patterns in written communication with cognition [473]. Borrowing from prior work [206], I adopt the following techniques to examine cognitive changes through linguistic syntax, structure, and stylistic vocabulary usage:

Coleman-Liau Index (CLI) is a measure of linguistic structure and provides a readability assessment based on character and word structure within a sentence [482]. This measure approximates a U.S. grade level required to understand the content, and can be calculated with the formula: $CLI = 0.0588L - 0.296S - 15.8$, where L is the average number of letters per 100 words and S equals the average number of sentences per 100 words.

Complexity and Repeatability are syntactic measures that indicate an individual's cognitive state in the form of planning, execution, and memory, and are in turn, linked to psychological states [206]. I quantify complexity as the average length of words per sentence, and repeatability as the normalized occurrence of non-unique words.

LIWC. I use LIWC [473] and specifically focus on the normalized occurrences of *Cognition & Perception*, *Linguistic Style*, and *Social Context* categories.

3.3.2 Results

I start with an overview comparing the differences between the changes in *Before* and *After* samples per dataset, CR, B_1, and B_2 in Table 3.10. To evaluate statistical significance of these differences, I conduct Welch's t-test, and adjust the p-values using False Discovery Rate (FDR) correction. For most of the measures, the *Treatment* and *Control* groups in B_1 and B_2 show no statistically significant differences in the *Before* and *After* periods, but all other measures barring one (*Activity*) show significant differences in the *Treatment* and *Control* groups in the CR dataset. This dataset also shows significant changes in magnitude for the *Treatment* group, for example – a) for *affect*, grief expression significantly increases, b) for *behavior*, increased social engagement, interactiveness, and diversity of interests, and c) for *cognition*, improved cognitive and linguistic processing.

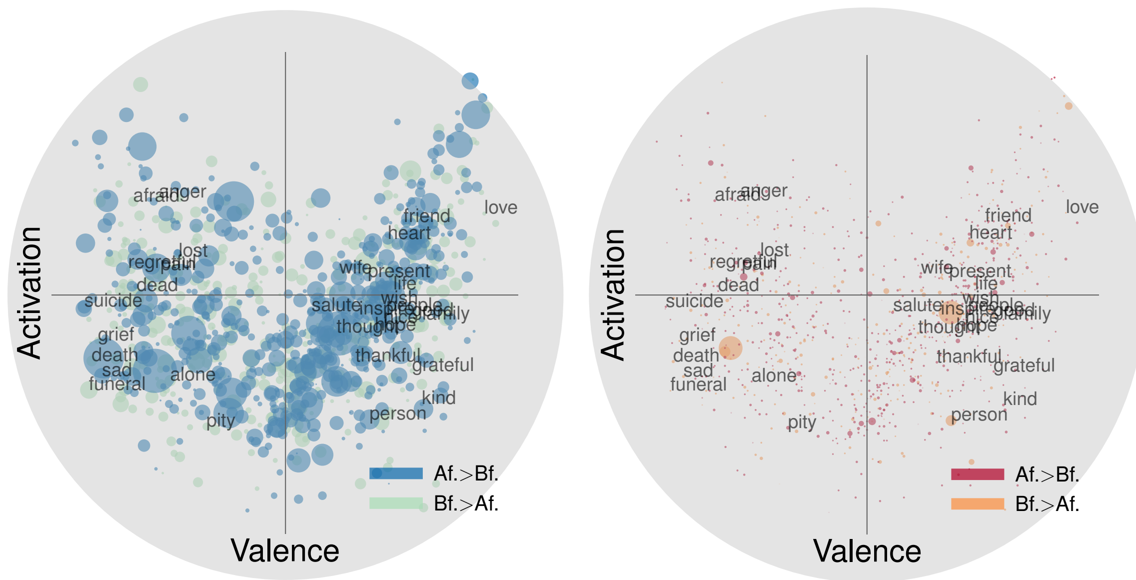
Several studies in psychology and the crisis literature have associated greater expressivity whether in terms of the positivity or intensity of emotionality, bereavement and grief expression, or language with an improvement in their psychological wellbeing status [347]. I situate this study's results within these studies to observe that compared to baseline scenarios, counseling recommendations following student deaths are succeeded by effects indicative of improved wellbeing.

Table 3.10: Comparing the mean percentage difference between *Before* and *After* periods in the Treatment (Tr) and Control (Ct) groups. Bar lengths represent relative and numbers denote absolute magnitudes. Blank entries convey no statistical significance.

Data →	**CR**		**B1**		**B2**	
Measure ↓	**ΔTr**	**ΔCt**	**ΔTr**	**ΔCt**	**ΔTr**	**ΔCt**
Affective Changes						
Grief: Activation	15	-1	–	–	–	–
Grief: Valence	9	-1	–	–	–	–
Behavioral Changes						
Activity	–	–	–	–	–	–
Interaction Diversity	9	8	34	27	–	–
Interactivity	29	-1	–	–	–	–
Cognitive Changes						
Readability	14	11	3	-1	11	11
Complexity	1.3	.7	5	6	.4	.6
Repeatability	-3	9	1	1.5	.5	3
Linguistic Style	481	92	–	–	–	–
Cognition & Perception	457	70	–	–	–	–
Social Context	382	49	–	–	–	–

Affective Changes

I examine the affective changes that characterize the *Treatment* group's exposure to counseling recommendation. I employ the circumplex representation of grief words, to find that grief expressions considerably increase (15% for valence, 9% for activation) in *Treatment* as compared to a marginal (-1%) decrease in *Control* ($t = 2.68, p < 0.05$). Figure 3.12 plots these changes from the *Before* to the *After* period on the same circumplex model, where larger circles indicate greater differences for those corresponding grief expression. A closer look at Figure 3.12(a) reveals that higher differences are more prominent in the cases where a specific grief expression increased in the *After* period. These expressions which show significant changes, belong to all the four quadrants in the circumplex model, such as "friend", "hope", "sad", and "lost". In contrast, although drawn on the same scale, large circles are scarce in Figure 3.12(b), suggesting minimal changes in grief expression in the *Control* group. This observation affirms that individuals exposed to counseling recommendations in the CR dataset become more expressive from an affective perspective, and this affective expression illustrates grieving as a positive psychological response to the crisis (that is, the



(a) Treatment

(b) Control

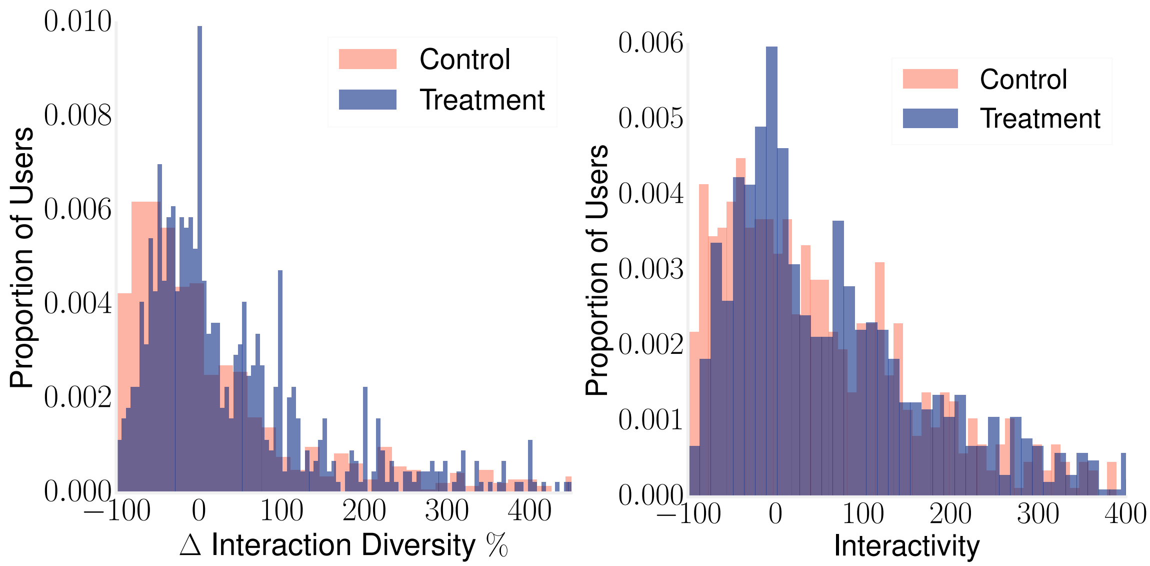
Figure 3.12: Differences in grief words (from the proposed grief lexicon) in the *Treatment* and *Control* groups, plotted on Russel's circumplex model of affect. The radius of the circles are proportional to the mean differences in occurrences of the grief words between the *Before* and *After* periods around the date of CR post.

student death incidents) [473].

Behavioral Changes

The findings suggest that counseling recommendations are associated with no significant differences in terms of a user's posting frequency (activity). An alternative interpretation of this finding backs our causal analysis that, despite all users continuing usual social media activity before and after the exposure to the CR post, the outcome varies for the *Treatment* and *Control* groups for "every" other measure.

Next, Figure 3.13 shows the behavioral changes in users around the date of sharing of the CR posts. For interaction diversity, that is, the measure of a user's engagement across multiple communities, we find similar changes in the *Treatment* and *Control* group, the former being marginally higher by 1% ($t = 4.0, p < 0.05$). However for interactivity, a major increase by 29% occurs in the *Treatment* cohort, as compared to a small -1% change



(a) Interaction Diversity

(b) Interactivity

Figure 3.13: Distribution of differences of interactivity (comments to posts ratio) and interaction diversity (unique subreddits).

in *Control* ($t = 4.1, p < 0.05$). These measures support positive social functioning effects of CR posts, in turn known to have coping benefits following loss of someone close [473].

Cognitive Changes

Readability Within the *Treatment* group in the CR dataset, Coleman-Liau Index (CLI) shows a mean increase of 14% in the following exposure to the counseling recommendations. Although this number is close to the changes in *Control* group (11%), there statistically significant differences ($t = -81,\ p < 0.05$) between the two groups. Since both groups of users were statistically matched on their overall linguistic usage, and are alike in their educational qualification (college students), a comparable overall increase in readability is unsurprising, especially because this measure typically increases with writing over the years [482]. To illustrate this observation further, I obtain the probability density function (with Gaussian kernel) of CLI in the *Before* and *After* periods of exposure to CR posts, for the *Treatment* and *Control* cohorts (Figure 3.14). This figure shows that the distribution of the CLI measure changed considerably for the *Treatment* group, and no such effect is observable in the *Control* group. Specifically, the variance of distribution in *Treatment*

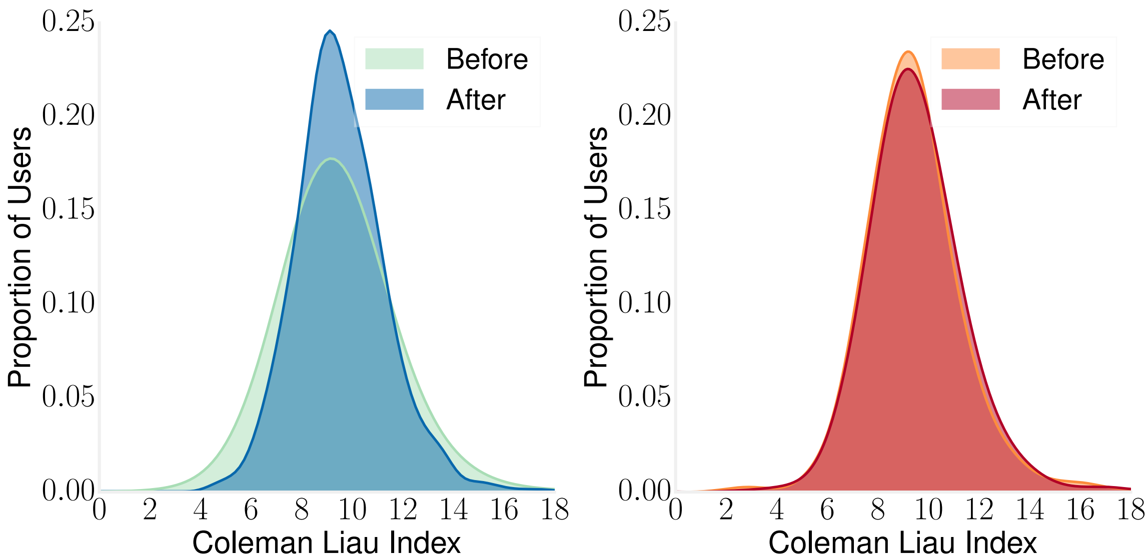
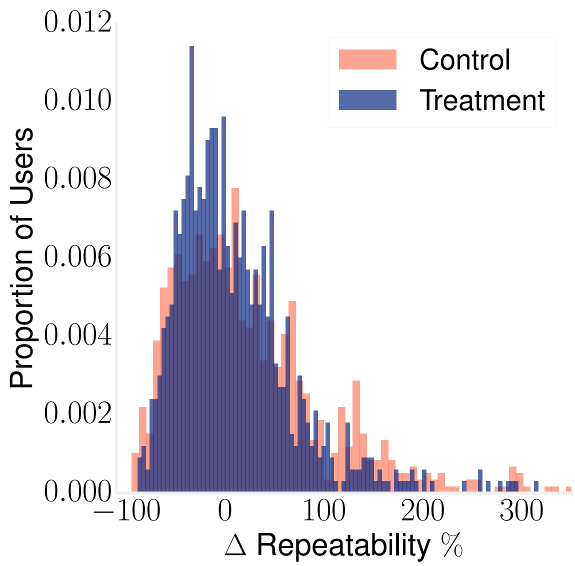


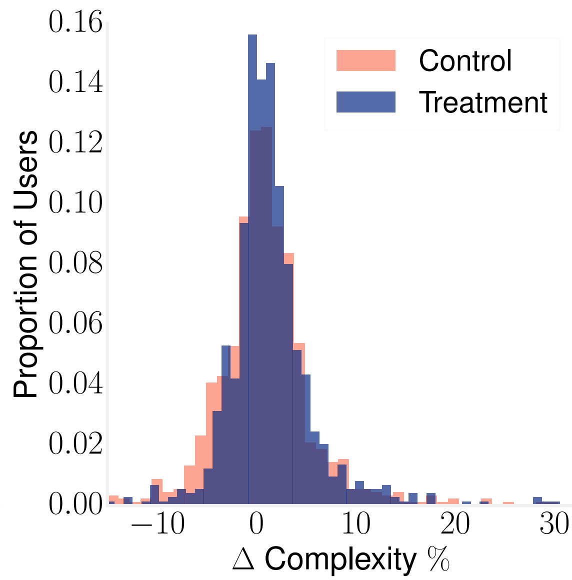
Figure 3.14: Distribution of Readability (CLI) in the *Treatment* (left) and *Control* groups (right), *Before* and *After* the CR post.

cohort reduced substantially by 90% (σ decreased from 6.1 to 1.9) after CR post exposure. Increased readability of written speech is known to indicate better control over the train of thought, better coherence in expressing ideas, and better discourse organization [618]. That such increases manifest in the *Treatment* group after exposure to CR posts further indicate psychological effects around improved wellbeing.

Repeatability and Complexity Figure 3.15 shows the *After* and *Before* differences in linguistic repeatability and complexity in the *Treatment* and *Control* groups following exposure to CR posts. For repeatability, the figure reveals that a greater fraction of *Treatment* users show negative and near-zero changes ($Mdn_{\text{Treatment}} = -2$ vs. $Mdn_{\text{Control}} = 8$), that is, their linguistic repeatability decreases. In addition to statistically significant differences ($t = 11.3,\ p < 0.05$), while repeatability decreases by 3% for *Treatment* users, it increases by 9% for *Control* users. For complexity, *Treatment* users demonstrate over 80% increase compared to the *Control* users (1.3% vs. 0.7%). Although numerically the change is small, statistical significance tests ($t = 18.6,\ p < 0.05$) show compared to a linguistically matched *Control* population, the *Treatment* users show a greater increase in the usage of longer words.



(a) Repeatability



(b) Complexity

Figure 3.15: Distribution of differences in repeatability and complexity in *Treatment* and *Control* groups.

Mental health challenges can manifest in the form of poverty of speech, are accompanied by a reduction in syntactic complexity, and an impairment in syntactic comprehension [206]. Such tendencies typically result from an overall cognitive deficit, difficulty concentrating, distraction, or a preference for expressing simpler ideas. As repeatability and complexity capture such syntactic attributes in Reddit posts, reduction in repeatability and increase in complexity following CR post exposure are, therefore, indicative of positive psychological changes in the *Treatment* cohort.

Cognition and Perception, Linguistic Style, Social Context. Finally, I analyze the normalized occurrences of LIWC categories for linguistic style, cognition, and social context. Figure 3.16 shows the variability (95% confidence interval) of differences for statistically significant LIWC categories. All of the categories in the *Treatment* dataset shows significantly higher variability than the *Control*. These plots lie on the positive y-axis, suggesting that the levels of cognitive measures increased following exposure to the CR posts.

The *cognitive* measures, such as "causation", "cognitive mechanics" and "tentativeness"

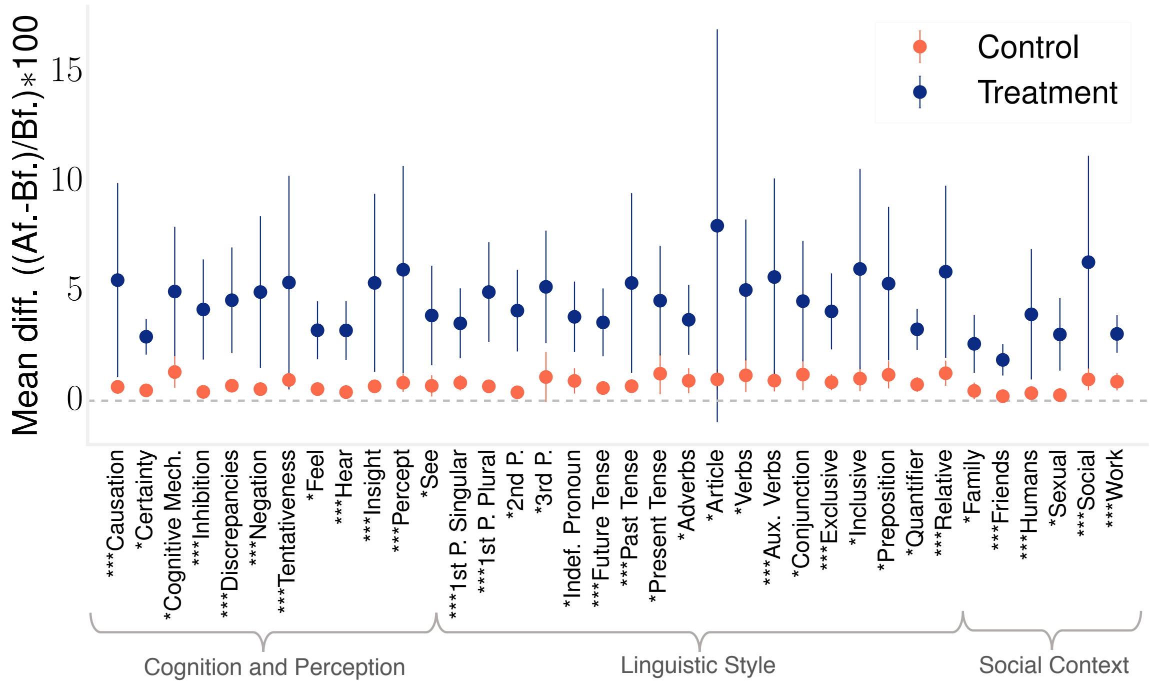


Figure 3.16: Differences in cognitive measures between the *Treatment* and *Control* groups following CR exposure, based on usage of LIWC categories. The vertical lines represents 95% confidence interval range, and the dot shows the mean. Statistical significance is reported based on Welch t-test. p-values are adjusted using FDR correction ($*p < 0.05, **0.001 < p < 0.01, ***p < 0.001$).

significantly increase after the exposure to CR posts. Per prior work, this indicates an improvement in an individual's cognitive functioning [126, 473]. Additionally, greater usage of "negation", and words relating to "feel" and "percept" indicate greater perceptual expressiveness, known to be associated with first-hand accounts of real world happenings, events, and experiences [82].

Likewise, within *linguistic style* measures, pronouns (1^{st}, 2^{nd}, and 3^{rd}) and temporal attributes considerably increase (mean difference=∼5) in the *Treatment* dataset. Both psycholinguistics and crisis literature note that 1^{st} person and past tense usage relate with narrating personal or collective experiences of upheavals, which seems likely in our case [397]. Prior work also notes higher usage of 2^{nd} person pronouns in the aftermath of crises and 3^{rd} person pronoun use is associated with the language of adaptive and coping related health benefits following crises. Further, the increased usage of lexical density features such as "adverbs", "articles", and "quantifiers" indicate that *Treatment* users express via more complex narratives [126]—a signal of better psychosocial health [206]. Among the *social context* measures, treated users use more "family" and "friends" words. Based on prior work, this is a known behavior for individuals coping with grief and trauma, and reference to socialization has therapeutic benefits for an individual's psychological state [567].

3.3.3 Discussion

This study demonstrates that, with a novel causal framework and unobtrusively gathered social media data, it is possible to quantify, to what extent exposure to counseling recommendations following a student death on a college campus positively impacts an individual's psychological state. Therefore, this study bears the potential to complement existing techniques of assessing the effectiveness of intervention measures deployed after crises. In this way, this study advances the growing body of research in social media and health, opening up new avenues of addressing health challenges by employing social media as a mechanism of supportive mental health and crisis intervention delivery.

The findings suggest statistically significant psychosocial (affective, behavioral, cognitive) effects of exposure to counseling recommendations on the treated population as compared to a statistically matched control cohort. In assessing these psychosocial effects, the causal inference framework accounts for behavioral and linguistic covariates across the treatment and control groups, also eliminating confounds due to temporal variability in their Reddit activity. Further, this study compares against other baseline scenarios to reveal that the observed effects were characteristic of the specific context of student death related crises, instead of other latent factors.

A contribution of this study is a *grief lexicon* and a transfer learning based methodology to build it. Drawing on recent advances in computational linguistics research, this study expands a validated affect dictionary with word embeddings and employed it on public social media data. This technique can be used in other social media and health research that involves extracting domain-specific information, but where ground-truth data is limited and unlabeled data is plenty.

The findings provide support for the "grief work hypothesis" [562], that situates grief counseling and therapy as a way for working through loss. The treatment group shows greater affective expressivity of grief, greater desire for social connectedness and diversity in interactions, improved cognitive and perceptual processing, and emergent linguistic and stylistic complexity. Based on psychology and crisis literature around the healing and coping benefits of grieving [325], the findings indicate that exposure to counseling recommendations on social media after crisis events, signals effects associated with positive benefits for one's psychological state. This provides positive empirical evidence about the efficacy of post-crisis counseling recommendations on college campuses.

These findings are not only useful in helping gauge whether sharing counseling recommendations on social media are at all effective, but also can support crisis rehabilitation efforts on college campuses. Campus officials can utilize the outcomes of this study as a way to identify individuals who are not benefiting from these counseling recommendations. This can help them employ other proactive intervention measures to support their mental health. Broadly, this study can inform campus policy decisions around mental health outreach. This study also sheds light into the role of communication technologies like social media, in supporting these efforts, both during crises as well as to tackle college student mental health challenges [196].

CHAPTER 4

SOCIAL MEDIA STUDIES OF WELLBEING IN WORKPLACES

This chapter discusses social media based modeling approaches of assessing wellbeing in another form of situated community — workplaces. Employee satisfaction and wellbeing is of prime interest to both individuals as well as organizations. Researchers have attributed employee subjective wellbeing as one of the prime determinants of important outcomes that range across, 1) health and longevity, 2) income, productivity, and organizational behavior, and 3) individual and social behavior [170].

I conduct two studies motivated address the gaps in state-of-the-art assessments of workplace wellbeing metrics. These studies bear implications in designing individual- and organization- facing tools to improve organizational functioning and wellbeing.

In the first study, we empirically study organizational culture by leveraging large-scale employee-contributed workplace experiences posted on Glassdoor. We examine the linguistic dynamics in public-facing anonymized reviews to describe culture, and develop a theoretically-grounded rendition of organizational culture as a codebook. We develop a lexicon to encapsulate culture based on 41 dimensions, and model organizational culture for company sectors and test its explanatory power in predicting employee performance, where we found that our computational model of organizational culture significantly explains individual performance and citizenship behavior, beyond individual intrinsic attributes (eg., demographics and personality). This work bears implications in designing individual- and organization- facing tools to improve organizational functioning.

In the second study, we quantitatively estimate role ambiguity via LinkedIn data. We compute LinkedIn based Role Ambiguity (`LibRA`) as a difference in one's self-described roles (on LinkedIn) and the company-published job description of the same role. We measure these differences using word-embeddings on the multiple dimensions of job aspects.

Aligning with a set of theory-driven hypotheses, we find that greater `LibRA` is associated with depleted wellbeing, such as increased heart rate, increased arousal, decreased sleep, and higher stress. In addition, `LibRA` is associated with lower job performance such as decreased organizational citizenship behavior and decreased task performance. We explored the self-presentation behavior and social computing platform-specific nuances and factors that need to be accounted if measures like `LibRA` are to be used in practice.

4.1 Social Media and Multimodal Sensing Data of Workplace Employees

Before we go into the details, we describe participant pool and the dataset of the two studies. The dataset primarily comes from a large-scale multi-sensor study of workplace behaviors, called the Tesserae Project [406, 422, 528]. This study, approved by the Institutional Review Board (IRB) at the researchers' institutions, recruited 757 participants, who are information workers in cognitively demanding fields (e.g. engineers, consultants, managers) across the United States. These participants who were recruited from January 2018 through July 2018, completed an initial set of questionnaires related to demographics, job performance, personality, intelligence, affect, anxiety, alcohol and tobacco use, exercise, sleep, and stress, personal attributes, and wellbeing, administrated via psychometrically validated survey instruments, as well as received daily surveys on a set of these attributes. Participants also received three sensors: location-tracking Bluetooth beacons; 2) a wearable; 3) a phone agent—a smartphone application [651]. In addition, some participants authorized collection of their historical social media data. As compensation, participants either received a series of staggered stipends totaling up to $750 or they participated in a set of weekly lottery drawings (multiples of $250 drawings) depending on their employer restrictions. Because the participants were enrolled over a 6 month period of time (January to July 2018) in a staggered fashion, data collection varied with a range of time between 59 days and 97 days (68 days on an average).

Participant Privacy and Consent. Given the sensitive nature of the data being collected, participant privacy was a key concern in the study. The participants were provided with an informed-consent describing each sensing stream, and technical specifications listed what each device was capturing and how it would be secured and stored. The participants needed to consent to each sensing stream individually, and they had the provision to clarify their queries / concerns about the sensing streams, and they could opt out of any of them [406]. Their data was de-identified and stored in secured databases and servers which were physically located in one of the researcher institutions, and had limited access privileges. Participants were made aware that they could voluntarily drop out via an email at any point during the year-long study period. Participants could also specifically request their data deletion from the database.

4.1.1 Self-Reported Data

As mentioned above, the enrollment process consisted of responding to a set of initial survey questionnaires related to demographics (age, sex, education, type of occupation, role in the company, and income). Participants were additionally required to answer an initial ground-truth battery, a set of survey questionnaires that measured their self-reported assessments of personality traits and executive function. Throughout their study period, they received daily or periodic validated surveys that recorded their self-reported assessments of job performance.

Related to psychological traits of the individuals, we collected, 1) *Cognitive Ability* (or executive function), as assessed by the Shipley scales of Abstraction (fluid intelligence) and vocabulary (crystallized intelligence) [578], 2) *Personality Traits*, the big-five personality traits as assessed by the Big Five Inventory (BFI-2) scale [589, 614], and 3) *Wellbeing*, the general positive and negative affect levels as assessed through the Positive And Negative Affect (PANAS-X) scale [658], the anxiety level as measured via State Trait Anxiety Inventory (STAI-Trait scale) [591], and the quality of sleep as measured via the Pittsburg

Sleep Quality Index (PSQI) scale [182].

Relevant to job performance, we describe two measures that we collected:

Task Performance To assess task performance, we use two scales, IRB (In-Role-Behavior) [663] and ITP (Individual Task Proficiency) [264]. The IRB scale contains seven items including questions such as *adequately performed assigned duties, failed to perform essential duties, performed expected tasks,* etc., each of which can be rated on a scale of 1 (strongly disagree) to 7 (strongly agree). On the other hand, the ITP scale contains three items, *carried out core parts of the job well, completed core tasks well using standard procedures, and ensured that the tasks were completed properly*, each of which can be rated on a scale of 1 (very little) to 5 (a great deal). Together, these instruments measure an individual's ability to adequately execute their assigned duties, and their proficiency at performing activities that drive an organization's technical core [70, 643].

Organizational Citizenship Behavior . We administer the OCB scale to measure organizational citizenship behavior [226]. Organizational citizenship behaviors characterize an individuals activities that are not typically or formally rewarded by the management, or voluntary activities outside one's core responsibilities, but which promote the welfare and effectiveness of the organization and its members [139, 455]. The survey instrument contains eight items, each of which asks the participant to self-reflect (yes/no), if they, *went out of their way to be a good employee, were respectful of other people's needs, displayed loyalty to my organization, praised or encouraged someone*, etc.

Passive Sensing Data for Offline/Physical Activity

To passively sense participants' behavior and wellbeing measures of participants, the study deployed three modalities of sensing technologies, as briefly described below.

Bluetooth Beacons. Participants were provided with two static and two portable Bluetooth beacons (Gimbal [238]). Static beacons were to be placed at their work and home locations, and the portable beacons were to be carried either in their backpacks or their wallets. Combined, these beacons tracked participant presence at home and/or work location, and also help to assess their commute and time at their work desk.

Wearable. Participants were provided with a fitness band based smartwatch (Garmin Vivosmart [234]), which they wore throughout the day. The wearable continually tracked and recorded health measures, such as heart rate variability, stress, and physical activity in the form of sleep, footsteps, and calories lost.

Smartphone Application. A smartphone application [423, 651] was installed on Participant smartphones (Android and iPhones) . This application tracked phone use, lock or unlock behavior, call durations, and also leveraged their smartphone-based mobile sensors to track their location (mobility), and physical activity.

4.1.2 Social Media Data

Social media was deployed as a passive sensing modality of behaviors and wellbeing in Tesserae [406, 528]. The study asked the participants to provide their Facebook and LinkedIn data, *unless they did not consent to do so, or did not have either account*, that is consent was sought only from those participants who had existing Facebook or LinkedIn accounts from before the study. They could optionally consent to their Instagram, Twitter, GMail (metadata only), and Google Calendar data. To collect the social media data from those who consented, we hosted a Python based web application that was developed by our team. This web application was built upon the Django framework and used an Open Authorization (OAuth) based data collection strategy. Compared to other alternative data collection strategies such as downloading and sharing of social media archives, or scraping through webpage crawlers or smartphone applications, the OAuth protocol provides a more privacy-preserving, streamlined, and convenient means of data collection at scale. Additionally, this not only poses minimal burden to the participants, but also ensures data sharing over a secured channel without transfer of any personal credentials. The following paragraphs explain our data collection infrastructure.

4.1.3 Privacy and Ethics

The Tesserae project was approved by the Institutional Review Board at the researchers' institutions. Given the sensitivity of the data, participant privacy was a key concern. The participants were provided with informed-consent documents describing the specifics of what data they were providing, and how would that be stored. The participants needed to consent to each form of data, and could also clarify concerns and opt out of any data collection. The data was de-identified and stored in secured databases and servers which were physically located in the researcher institutions, and had limited access privileges.

4.2 Modeling Organizational Culture with Crowd-Contributed Workplace Experiences

In organizations, certain norms and principles that are believed to optimize the workforce and maximize efficiency are referred to as *organizational culture* (*OC*) [39, 445]. This embodies a core value system which affects the development and execution of new ideas, and the management of unexpected events like crises [107, 444]. While metrics such as revenue and profit are standard methods to gauge the effectiveness of an organization, the culture of an organization is both an indicator and a factor to influence its effectiveness [601]. From an employee's perspective, comprehending organizational culture can help foretell their loyalty and commitment [444] because community can affect human behavior [102, 149].

Organizational studies have employed a variety of survey instruments to quantify *OC*, but these come with their own challenges [133, 135, 240, 304, 495]. These instruments lack temporal granularity and do not scale. Besides, conducting such studies in organizational settings leads to unique problems because of employee anxieties regarding the confidentiality of their opinions [30, 431]. Therefore, the workplace context can invite multiple biases, such as response (or non-response) bias, study demand characteristics, and social desirability bias [45].

In contrast, workplace review platforms contain self-initiated and anonymous reports [242] that stand to mitigate many of the biases introduced by survey studies [244]. Glassdoor is one such platform with publicly posted reviews of workplace experiences. Not only do these reviews contain objective information like pay, hours and benefits but also the free-form text that encapsulates various nuances of *OC* [39, 276]. Through the affordance of descriptive text, platforms like Glassdoor provide an accessible, scalable and flexible medium to express cultural and ecological differences [250]. Our work leverages the language used in publicly visible employee reviews to computationally model *OC* and augment our understanding of it. Specifically, this study has the following research aims:

Aim 1. To operationalize *OC* as a multi-dimensional construct and validate it with language on Glassdoor.

Aim 2. To computationally model *OC* of an organizational sector, and evaluate if it explains employee job performance.

The first research aim strives to build a usable construct of *OC*, based on Glassdoor data, that captures various aspects like interpersonal relationships, work values, and structural job characteristics. Towards this, I use established frameworks from the domain of organizational psychology [133, 135, 240, 304, 495] to identify job descriptors related to *OC* and represent them as word-vectors. Grounded in the literature, this study models organizational culture in the lexico-semantic space of word embeddings [476], and validates this word embedding based construct of *OC*. This produces a codebook of lexical phrases that closely align with different dimensions of *OC*.

Next, given a reliable representation of *OC* I examine if it explains individual performance [444, 445] by quantifying the *OC* of companies by sector (e.g., management, production, or computer). On a ground truth dataset from the Tesserae project [406, 528] with 341 employees from three companies this study finds that incorporating *OC* improves on intrinsic traits (such as demographics and personality) to explain an employee's task performance and citizenship behavior. This renders empirical evidence that *OC* explains human functioning and exhibits an application of our construct.

4.2.1 Data: Glassdoor

Glassdoor as Employee Experience Platform

For this study, crowd-contributed workplace experiences from Glassdoor serve to validate the operationalized *OC* (Aim 1), and to quantify the *OC* in an employee sector (Aim 2). Glassdoor is an online platform (launched in 2008), for current and former employees to write reviews about their workplace experience. As of 2018, there are 57M individual accounts on this platform, and there are 35M reviews posted for 770K companies [243]. Salient topics in these reviews include work-life balance, management, pay, benefits, growth opportunities, and colleagues. Glassdoor reviews require ratings and free-form text. Employees can rate their overall experience on a scale of 1 to 5, and optionally add ratings for fields like career opportunities, compensation, and senior management. The free-form text field requires employees to submit descriptions of their workplace experience, in separate sections for *Pros* and *Cons* (Table 4.3). This text describes many salient workplace themes, such as work-life balance, management, pay, benefits, growth opportunities, facilities, and interpersonal relationships.

Quality of the Content

In Glassdoor's published community guidelines and norms for content submission, they state that they *strive to be the most trusted and transparent place for today's candidate to search for jobs and research companies* [244]. Both contributing content and consuming content

Table 4.1: 41 Org. descriptors from O*Net to represent the dimensions of *OC*. The category column indicates the O*Net category of the descriptors. Categories with '*' are subcategories within the "Work Context" cateogry. The table in supplementary material provides a detailed description of job descriptors with the validation source.

Category	Organizational Culture Dimensions
Interests	Conventional, Enterprising, Social
Work Values	Relationships, Support, Achievement, Independence, Recognition, Working Conditions
Wk. Activities	Assisting & Caring for Others, Establishing & Maintaining Relationships, Guiding & Motivating Subordinates, Monitoring & Controlling Resources, Training & Teaching Others, Coaching & Developing Others, Developing & Building Teams, Resolving Conflicts & Negotiating
Social Skills	Instructing, Service Orientation
Struct. Job Characteristics*	Consequence of Error, Importance of Being Exact, Level of Competition, Work Schedules, Frequency of Decision Making, Freedom to Make Decisions, Structured versus Unstructured Work
Work Styles	Concern for Others, Leadership, Social Orientation, Independence, Integrity, Stress Tolerance, Self Control, Adaptability, Cooperation, Initiative, Achievement
Interpersonal Relationships*	Frequency of Conflict Situations, Face-to-Face Discussions, Responsibility for Outcomes & Results, Work w. Group or Team

necessitates an individual login. It only allows individual accounts with *permanent, active email address, or a valid social networking account* to submit content, with a maximum allowance of *one review, per employee, per year, per review type* [245]. Glassdoor moderation involves proprietary content-analysis technology as well as human moderators. Any reviews deemed to be incentivized or coerced, are either not allowed or removed from the platform. In addition, Glassdoor offers the option to flag content, which is evaluated on a case-by-case basis. To ensure a non-polarized distribution of reviews, Glassdoor implements a key incentive policy known as, "give to get" [242]. In this model to get full access to all reviews, viewers must contribute their own review. This paradigm encourages more neutral opinions to be recorded and diminishes the biases of self-selected users [108]. The content posted on Glassdoor remains anonymous, and the moderation strategies ensure that no sort of individual-identifiable detail is disclosed in the content. However, each review comes tagged with the reviewer's role, employment status (current or former), and location of employment.

4.2.2 Aim 1: Operationalizing Organizational Culture

To measure *OC* through language on Glassdoor reviews, this study first operationalizes it based on language. I adopt a three-step approach to achieve this: 1) identifying descriptions of multiple dimensions of *OC*. 2) transforming the descriptions into word-vectors to capture their linguistic and semantic context, so as to represent *OC* as a collection of these vectors. and 3) comparing the word-vector based *OC* construct to filter Glassdoor posts related to *OC* and qualitatively investigate the posts' keywords to establish face-validity.

Identifying Descriptors of Organizational Culture

Language used by a community (or organization) provides a unique lens to interpret its culture [39, 250]. To understand the extent to which a text expresses *OC*, I needed an established ontology of job aspects that are indicative of different *OC* dimensions. For this, I obtain job aspect descriptors from the Occupational Information Network (O*Net). O*Net (*onetonline.org*) is an online database of occupational information developed under the sponsorship of the U.S. Department of Labor/Employment and Training Administration (USDOL/ETA).

These descriptors are motivated by organizational research [257, 293, 610], and are regularly updated with changes in socio-economical and workforce dynamics. O*Net describes 189 different job descriptors, categorized in 17 sub-categories, which are further grouped into 8 primary categories. Each of the 189 job descriptors, like *Stress Tolerance*, *Level of Competition* and *Independence*, is accompanied by a description.

However, all descriptors do not necessarily explain *OC*. For example, *Staffing Organizational Units* and *Pace Determined by Speed of Equipment* simply describe characteristics of the job role, not the underlying concept of *OC*. Therefore I verify which descriptors align with established frameworks of *OC* that are widely used in organization research. Two coauthors familiar with organizational studies independently inspected each of the 189 descriptors in O*Net on the basis of four *OC* instruments, *Organization Cultural Inventory* [134]), *Organization Culture Profile* [445]), *Hofstede's Organization Culture Questionnaire* [304], and *Organization Culture Survey* [240]). Any discrepancies ($n = 23$) with respect to the validity of a job descriptor was resolved by both authors on agreeable themes and concepts. Overall this procedure had a Cohen's κ (inter-rater reliability) score of 0.89 This process retains 41 descriptors, each of which describes an aspect of *OC* (see Table 4.1). Also note that these dimensions are not necessarily mutually exclusive or disjoint [445, 520], and we could expect a significant overlap in our ensuing analysis. Our domain-driven approach validates the O*Net descriptors on the basis of multiple different frameworks because no single conceptual framework describes *OC* exhaustively [133, 520].

Transforming Descriptors into an OC Construct

While O*Net provides explanations of the 41 descriptors of *OC*, simply tokenizing the keywords in these descriptions would not adequately capture the concept of *OC*. Therefore to address this challenge, I encapsulate the linguistic and semantic context of these descriptions by using the concept of word embeddings [211, 544]. This approach represents words as a vector in a higher dimensional space, where contextually similar words tend to have vectors that are closer.

I use pre-trained word embeddings in 50-dimensions (GloVe: trained on word–word co-occurrences in a Wikipedia corpus of 6B tokens [476]). Building on prior work of representing job aspects in lexico-semantic dimensions [535], I transform the explanations for each of the 41 descriptors (Table 4.1) into a 50-dimensional word-embedding vector. These 41 word-embedding vectors essentially characterize multiple dimensions of *OC* in a latent semantic space. Collectively, they constitute our operationalized construct of *OC*.

Validating the Operationalization of OC

While our operationalization of *OC* captures the information contained in 41 descriptors (obtained from O*Net and validated from domain assessments of *OC*), I need to establish its validity for practical use. The research team qualitatively inspects the top keywords in text from our Glassdoor dataset that is relevant to *OC*.

Table 4.2: Descriptive stats. of Glassdoor dataset of 92 companies (sourced from top 100 of Fortune 500). Aggregated values are per company.

Measure	Total	Mean	Stdev.
Reviews	616,605	6,702	8312
Pros Sntncs.	1,386,787	15,073.77	18,408.64
Pros Words	10,747,265	17.42	20.91
Cons Sntncs.	1,715,875	18,650.82	22,786.10
Cons Words	17,150,342	27.81	47.24

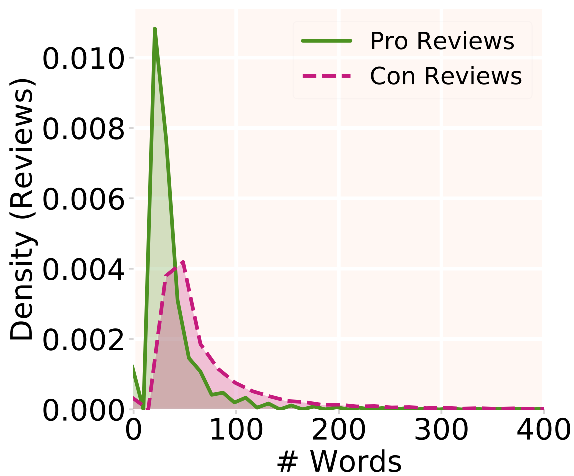


Figure 4.1: Distribution of number of words per review in the Glassdoor dataset of Fortune 100 companies.

Compiling the Glassdoor Dataset

To obtain a diverse but voluminous dataset on Glassdoor, I consult the *Fortune 500* list (ranked by revenue) [222] and obtain the top 100 ranked companies. Since only 8 of these companies appear in the list of *Fortune 100 Best Companies to Work For* [223], I believe the considered sample is not dominated by companies with positively-skewed employee experiences.

I obtain the public reviews of these organizations using web scraping. For each review, I collect the textual components (segregated into *Pros* and *Cons*) and the reviewer's employment information (role and location). Table 4.3 shows three example excerpts in *Pros* and *Cons* components. In sum, I obtain $616,605$ reviews from 92 companies (at the time of writing 8 companies did not have profiles on the platform) that were posted on Glassdoor

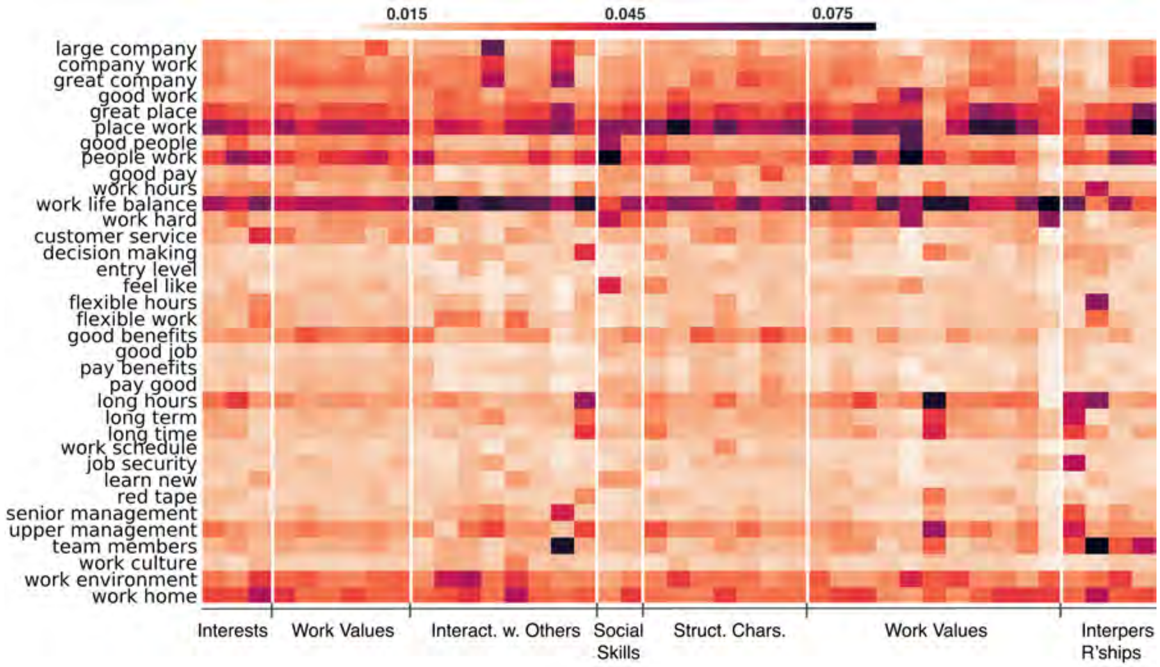


Figure 4.2: Top n-grams in sentences about *OC* (excluding lexical variants of keywords). Darker colors (higher TF-IDF score) indicate greater relative importance within a particular dimension. Dimensions have been categorized corresponding to the scheme in Table 4.1

Table 4.3: Example paraphrased excerpts in *Pros* and *Cons*.

Pros	Cons
1) Great teams 2) Talented co-workers 3) Not stressful 4) Good work-life balance	Most departments offer no flexibility in work schedule. My manager doesn't allow me breaks for doctor appointments, child's school activities
Good work environment, nice people. Lots of fun working on cool technology. Location is also superb.	No communication from upper management, Pay is not nearly as competitive as market salaries.
Friendly, outgoing coworkers. Very health-conscious environment. Activities are encouraged and supported.	Little recognition for overtime hours, no WFH alternatives even with bad weather, poor work-life balance

between February 20, 2008 and March 22, 2019, amounting to $10,747,265$ words in the *Pros* segment and $17,150,342$ words in the *Cons* segment (ref: Table 4.2 and Figure 4.1). Note that the content distribution is skewed towards the *Cons*, but this observation aligns with activity on other review platforms [318]. Despite the possibility that some of these reviews could be capricious and circumstantial, this work intends to leverage the ample volume of data and capture themes at an aggregated level. Additionally, all our computation normalizes data by volume.

Table 4.4: The word-vector representation of these sentences that show a cosine similarity of 0.90 or greater for the corresponding *OC* dimension. Note that the same sentence can reflect multiple dimensions.

Example Text	OC Dimension
Great training, really genuine and supportive colleagues, great ways to get involved with interest groups— Proposal writing, research for new industry areas, volunteer activities	Social
In many instances rank was invoked just to prove a point, rather that using data for the same.	Importance of Being Exact
The drive to succeed is key, however, it's not a cut throat competition - people are humble and people at all levels are interested and willing to develop those at the lower career levels.	Level of Competition
If you have a goal and willing to work on it, senior management will have a genuine interest in helping you succeed.	Coaching and Developing Others
A lot of emphasis is on firm activities making it difficult to build relationships as you can only meet coworkers on Fridays, if they do come.	Establishing and Maintaining Interpersonal Relationships
New recruits are immediately given responsibility, and can take complete charge of their career development.	Initiative
Lot of group work makes the work easier and more fun.	Independence

Filtering Posts about Organizational Culture

First, I derive a word-vector representation of every sentence in the 616,605 posts ($\sim$3M) from the Glassdoor dataset. I use cosine similarity to measure the similarity between each sentence's word-embedding representation and each of the 41 dimensions of *OC* [34, 536]. Higher cosine similarity indicates that the sentence is semantically similar, or "talks about" that particular dimension of *OC*. I retain any sentence that exhibits a similarity of more than 0.90 with any of the *OC* dimensions. Note that the same sentence may express an opinion about multiple classes; for example, a post reading "Some staff is able to negotiate to avail work from home at least one day per week" relates to *Work Styles: Social Orientation*, *Work Values: Relationships*, and *Work Values: Independence*. Table 4.4 enlists a few paraphrased examples.

Establishing Face and Construct Validity

Since the sentences that clear the threshold only relate to *OC* through the latent semantic space of word-embeddings, I also investigate the actual language used in the content. I obtain the top 100 keywords (n-grams, n=2,3,4) in all sentences (above the similarity threshold of 0.90). Then, I compute the TF-IDF score for these keywords across each of the 41 *OC* dimensions (similar to [544]). Essentially, this reflects the importance of each keyword in the sentences that refer to an aspect of *OC*. Figure 4.2 visualizes the relative importance of these keywords (the supplementary document provides a heatmap with all top 100 n-grams). This study draws upon the validity theory [451], to establish face and construct validity of contextualizing *OC* in Glassdoor data by qualitatively examining the importance of the keywords in the *OC* dimensions.

The most dominant keyword across several dimensions is **work life balance**, and its lexical variants like "life balance", "work life". This recurrence could be because notions of work–life balance has many facets (beyond work-family conflict) such as personal needs, social needs and team work [463]. For instance, this n-gram is important to the *Social* dimension of *OC* because it characterizes altruistic behaviors and aid of colleagues [303]. Similarly, dimensions like *Assisting and Caring for Others*, *Coaching and Developing Others*, and *Training and Teaching others*, inherently overlap with the team based aspect of "work life" [240, 495]. Socially supportive and inclusive workplaces tend to foster better work–life balance, these key-words co-occur with language referencing social and interpersonal dimensions, for example *"[Company] tries to ensure work life balance, whether it works is another story as everyone seems too dedicated."* and *"[Company] offers the best work life balance and true diversity among big firms"*.

Certain keywords are relatively more discriminatory between *OC* dimensions. For instance, **"good benefits"** is salient in reviews about dimensions like *Support* and *Recognition*. For employees, reward systems within companies garner reciprocal loyalty and increase the perceived organizational support [198]. For example in this post, "There is effective communication from senior management along with a good benefits package, cutting-edge technology, and a culture of integrity and innovation that provides a very satisfying environment.". Another such keyword is **"job security"**, which is most relevant to experiences that refer to the *Frequency of Conflict Situations* dimension. This draws from the fact that employees in workplaces that have high disagreements require more security and stability of employment [304]. Other examples of identifiable n-grams are **"flexible hours"** or **"flexible**

work". These keywords gain maximum importance in text associated with the *Face-to-Face Discussions* dimension. Prior research found that teams with fluid hours accommodate more interactions [97]. Similarly, the terms **"long hours"** and **"long time"** are important in texts related to the *Stress Tolerance*. Longer working hours not only causes fatigue but also increases an employee's exposure to work-related stressors [101, 329, 590], such as that expressed in, *"Client projects can require* ***long hours*** *on short notice, and the general environment can be very demanding and not forgiving."*

Apart from those discussed above, some of the n-grams correspond to the dimensions of *OC* more intuitively. For example, **"good people"** is most important in texts associated with *Resolving Conflict* (*Interacting with Others*), **"senior management"** is relevant to texts about *Frequency of Conflict Situations* (*Interpersonal Relationships*), and **"team members"** dominates experiences about the *Face-to-Face Discussion* dimension (*Interpersonal Relationships*). The evidence provided by this study indicates that the *OC* construct built from curated O*Net job aspect descriptors can capture the *OC*-related language used in Glassdoor reviews.

4.2.3 Aim 2: Modeling OC and examining its Relationship with Job Performance

Prior work in the domain states that organizational culture (*OC*) influences individual performance in the workplace [640, 675]. This motivates this study to apply our 41-D model *OC* on posts of an organizational community (such as occupational sector) to explain the job performance of employees belonging to the same group. Here, I describe a methodology to computationally model *OC* with the proposed construct. Then, I evaluate whether the proposed model can augment our understanding of employee-functioning beyond what is explained by individual differences.

Table 4.5: Summary of individual attributes for Aim 2.

Measure	Scale	Range	Mean	Stdev.	Distribution
Independent Variables					
Demographics					
Age		21-64	34.15	9.01	
Gender	Categorical: Male | Female				
Job characteristics					
Tenure	Ordinal: 10 values [<1Y,1Y,..>8Y]				
Supervisory Role	Categorical - IT | Non IT				
Personality Traits (BFI-2)					
Extraversion	1-5	1.67-4.91	3.42	0.68	
Agreeableness	1-5	2.08-4.91	3.85	0.54	
Conscientiousness	1-5	1.92-5.00	3.90	0.65	
Neuroticism	1-5	1.00-4.67	2.44	0.75	
Openness	1-5	1.17-4.91	3.79	0.60	
Executive Function (Shipley)					
Crystallized: Abs.	0-25	0-23	17.11	2.97	
Fluid: Voc.	0-40	0-40	33.06	3.93	
Dependent Variables					
Job Performance					
IRB	7-49	20-49	44.48	4.57	
OCB	20-100	32-100	56.20	10.28	

Compiling the Ground-truth Dataset

Towards the Aim 2, I use the groundtruth dataset from Tesserae (section 4.1), and focus on three major companies, C_1, C_2, and C_3. I collect the Glassdoor reviews of these three companies. Our groundtruth dataset comes from the Tesserae project [406, 423, 528]. This provides us the individual difference attributes and job performance of 341 information workers across 18 unique sectors in three companies C_1, C_2, and C_3 in the U.S. Table 4.5 summarizes the distribution of these measures across the 341 individuals in our groundtruth dataset. The individual attributes include demographic details such as age, gender, education, supervisory role (supervisor / non-supervisor), income, and their role in the organization. This dataset also contains information on personality traits and executive function, both of which are robust indicators of job performance. The Big Five Inventory (BFI-2) scale [589, 614] measures personality traits across the big five personality traits of openness, conscientiousness, extraversion, agreeableness, and neuroticism. The Shipley scale [578] measures the executive function in terms of fluid and crystallized intelligence. The dataset provides two job performance measures *the IRB scale* [663] (In-Role Behavior) and *the OCB scale* [226]

(Organizational Citizenship Behavior).

The dataset classifies participants into 18 unique sectors based on role. The top three sectors by participant count are "business and financial operations" (115), "computer and mathematical" (105) and "management" (50), but the dataset also features sectors like "office and administrative support" and "healthcare practitioner". This leads to 25 combinations of company and sector (eg. {C_1, Computer and Mathematical}, {C_2, Management and Consultancy}, etc.). Corresponding to the same companies (C_1, C_2, and C_3) and the same sectors, I obtain $23,791$ reviews on Glassdoor (22,794 for C_1, 574 for C_2, and 423 for C_3). At an average of 350 reviews per {*company, sector*} group. These reviews contain 1,654, 134, and 108 unique roles respectively that mapped to the 18 sectors. For this, I use a semantic similarity based approach using pre-trained word vectors (trained on 6B tokens on the entire Wikipedia corpus) [476], and next, two researchers manually verified the mapping, and edited the sector label wherever necessary.

Modeling and Quantifying OC by Organizational Sector

Since culture is a collectively experienced and manifested, experiences expressed by employees who share a common basis, such as a team, department or sector in an organization are considered together. Such an approach facilitates a robust and replicable mechanism to study *OC* both between and within organizations — as prior work investigated the phenomenon on varying levels of organizational granularity [185, 580]. This study is motivated by recent social media language analyses that use word embeddings [535, 536] to model *OC*.

First, I collate all the reviews posted per company sectors. Then, using word-embedding based cosine similarity, I obtain the similarities of every review sentences with each of the 41 *OC* dimensions. One cannot simply apply the similarity measure directly as certain posts could be talking about a dimension either positively or negatively. Consider the *Independence* dimension (in *Work Styles*), which refers to a culture that expects employees to be unsupervised and self-motivated. For some employees, such a culture can be favorable

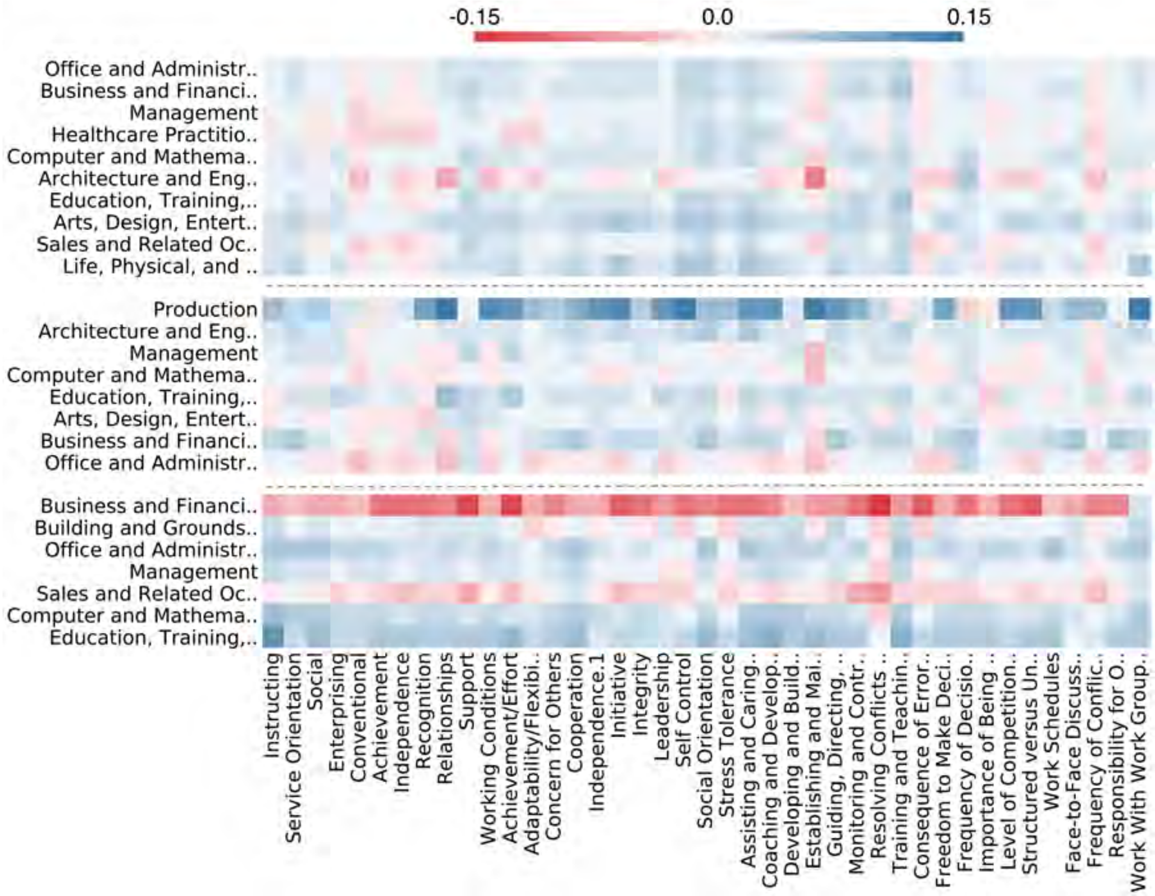


Figure 4.3: Organizational culture as quantified via Glassdoor data per organizational sector in three companies C_1 (top), C_2 (middle), and C_3 (bottom). The color and intensity of the cells represent the positivity or negativity in that dimension of organizational culture.

whereas for others it can be challenging. So, I qualify the raw similarity score between a post with the help of Glassdoor's *Pros* and *Cons* structure. I assign a weight of $+1$ to those sentences labeled as *Pros* and -1 to those sentences labeled as *Cons*. I obtain the weighted average of cosine similarities for each dimension. Together, this represents a 41-dimensional vector of *OC*, where a value per dimension is equivalent to how positive or negative that dimension is lexico-semantically spoken about in an organization's Glassdoor reviews. In this way, one can describe the *OC* of any group of employees in terms of a 41-D vector as long as one can retrieve a corpus of Glassdoor like experiences.

Figure 4.3 shows the distribution organizational culture in 41 *OC* dimensions in our dataset. *OC* varies across sectors both within and between companies. For example, the reviews from employees in the sector "business and financial operations" shows contrasting trends — while the reviews in C_1 and C_2 talk about *OC* in a similar way, the reviews of C_3 typically discuss dimensions of *OC* in *Cons*. Note that company characteristics of the scale and varying interests of employee-base could influence these sort of differences in the employee perspective on culture [171, 580].

Relationship between OC and Job Performance

As human behaviors are affected by the complex interplay between an individual and the culture they are embedded within [102], this study hypothesizes that its approach of operationalizing *OC* can explain an individual's job performance [444, 445, 675].

> **Hypothesis.** Organizational culture provides significant explanatory power towards one's job performance.

I test the hypothesis by predicting job performance — 1) In-Role Behavior (IRB) and 2) Organizational Citizenship Behavior (OCB). I build a baseline model (*Model 1*), with individual attributes, to predict job performance (Equation *Model 1*). This is motivated by prior work that extensively established that individual difference attributes (such as demographics, personality, and executive function) are strong indicators of job performance [42, 171, 261, 423, 496, 547, 557]. This study also controls for the individual's organizational sector. Next, this study builds an experimental model (*Model 2*), where I incorporate *OC* alongside the individual difference variables, and predict the same job performance measures again (Equation *Model 2*). Here, I include the 41-D representation of *OC* based on the Glassdoor posts of each employee's {*company, sector*}. If *Model 2* is better (statistically significant) in explaining the job performance measures than *Model 1*, then the hypothesis is supported.

$$JP \sim gender + age + income + supervisory\_role + tenure + exec.\_function + personality + org.\_sector \tag{Model 1}$$

Table 4.6: Summary statistics of the "best" regression models in *Model 1* and *Model 2*, where *Model 2* includes organizational culture, whereas *Model 1* does not. ***: p <0.0001

Measure	**IRB**		**OCB**	
	Model 1	*Model 2*	*Model 1*	*Model 2*
Algorithm	Lasso	Ridge	Ridge	Ridge
R^2	0.23***	0.28***	0.15***	0.24***
Pearson's r	0.43***	0.45***	0.32***	0.41***
SMAPE	3.67	3.65	6.96	6.71

$$JP \sim gender + age + income + supervisory\_role + tenure + exec.\_function + personality + org.\_sector + OC[41D] \qquad (\textit{Model 2})$$

Since the job performance measures are continuous, both models are regression estimators. I use three types of linear regression models with regularization, Lasso (L1 regularization), Ridge (L2 regularization), and Elastic Net (L1 and L2 regularization), and two non-linear regression models, SVM and Random Forest regressors. To tune the parameters of the models, I use grid search [585]. I use a leave-one-out (loo) [1] methodology to train and predict over our dataset, that is, this approach iteratively trains models with one held-out data sample, and predicts on that held-out sample. Finally, all the predicted data points are collated to obtain the pooled model performance measures — these include Pearson's correlation and Symmetric Mean Absolute Percentage Error (SMAPE) to evaluate the predictive accuracy of our models, and R^2 to evaluate the model fit (here JP is job performance).

Does Organizational Culture Explain Job Performance?

Model Performance Table 4.6 summarizes the fit and accuracy metrics of *Model 1* and *Model 2* for predicting job performance measures (IRB and OCB) (see Figure 4.4 for scatter plots). First, we observe that the data behaves linearly, as neither SVM regressors and Random Forest regressors performs better than the linear models. Next, and more

[1]The rationale to use loo validation over more standard k-fold cross-validation rests on the bias-variance tradeoff [641]. Given the small size of our dataset (n=341), such an approach leads to unbiased but high-variance models per fold. This ensures greater stability, robustness, and reduced randomness in sampling [352, 667].

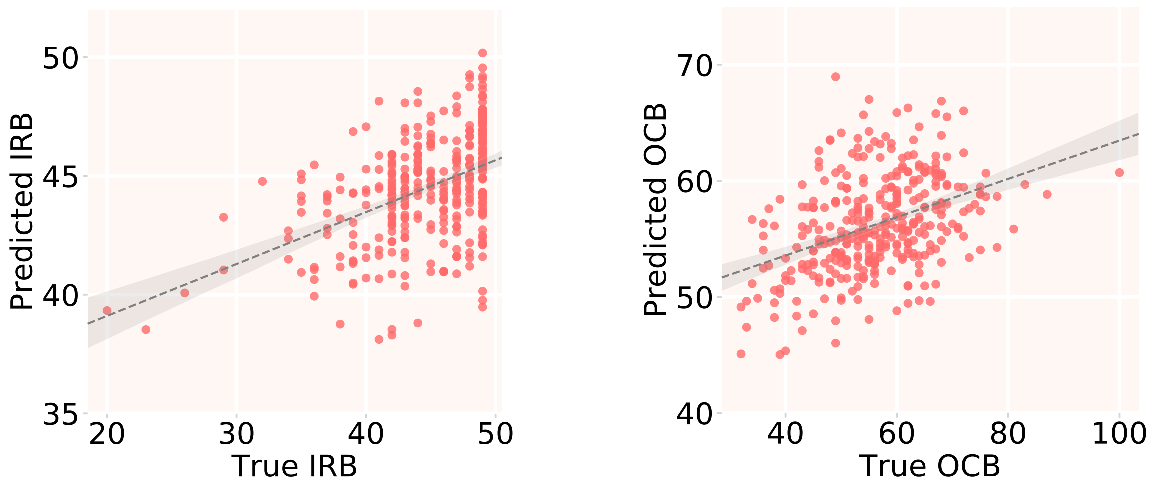


Figure 4.4: Scatter plots showing true and predicted values per *Model 2* of IRB (left) and OCB (right).

importantly, we find that *Model 2* which includes the organizational culture as an independent variable (or feature), performs better than the *Model 1*. In the case of IRB, for instance, *Model 2* fits 22% better, and *Model 2* predicts with 5% better-pooled correlation, and 0.6% lower SMAPE. In the case of OCB, the improvement is significantly high, with 60% better fit, 28% predicted correlation, and 4% lower predicted error compared to the performances on the job proficiency measures given by *Model 1*. All these models fit and predict with statistical significance ($p < 0.01$).

Model Validity. Despite *Model 2* performing better, it is important to reject the possibility that this is by chance. I aim to reject the null hypothesis that a randomly generated 41-D vector will perform better than our particular 41-D *OC* (*Model 2*). Drawing on permutation test approaches [17, 537], I run 10,000 permutations of randomly generated *OC* vectors. I find that the probability (p-value) of improvement by a randomly generated vector is 0.0002 for IRB and 0.0001 for OCB. This rejects the null hypothesis and reveals statistical significance in the observed improvement by including *OC* based on our quantification. Further, ANOVA tests to compare *Model 1* and *Model 2* reveals that *Model 2* fits significantly better (p<0.001) for both IRB (F=974) and OCB (F=310). Therefore, supporting our hypothesis, *OC* as computationally modeled using Glassdoor reviews, explains job performance of individuals.

Interpretation of Results

Table 4.7 reports coefficients of the top 10 *OC* dimensions (ranked on variable importance [266]) in *Model 2*. In-Role Behavior (IRB) assesses an employee's efficiency in accomplishing formal task objectives directly pertaining to their appointed job role. The positive relationship between *Recognition* and IRB is obvious because proficiency in one's assigned role leads to rewards through incentive and upward mobility [638]. *Responsibility for Outcomes and Results* is also positively related to IRB because individuals high in conscientiousness are known to be superior in task performance [32, 154, 372]. Experiences talking about *Frequency of Conflict* in the *Pros* more often correspond to higher IRB scores because conflicts (interpersonal, process-based or task-related) are detrimental to performance [326].

Organizational Citizenship Behaviors (OCBs) are not related to formal job roles and typically involve serving the community with extra-role tasks. The *Adaptability/Flexibility* dimension negatively associates with OCB because an *OC* which is more open to variable work styles triggers reduces face time between employees leading to fewer opportunities to give back [634]. This also explains why dimensions like *Work Schedule* and *Face to Face Discussions* are positively related to OCB. Also, OCB is based on mutual respect and reciprocality [142, 635]. This explains the positive relationship with experiences favorably describing the *Establishing and Maintaining Interpersonal Relationships* dimension. Additionally, work environments with high job autonomy elicit more OCBs as employees are empowered to use their time for altruistic outcomes [55]. The negative relationship with the *Conventional* dimension represents being clear of authority and rigidity.

Post-Hoc: Does Language tell us more than Ratings?

Finally, after establishing that quantifying *OC* with Glassdoor posts of a sector *does* significantly explain individual performance at workplace, I revisit the question, "is quantifying via language actually effective?" As Glassdoor is a platform that allows individuals to provide

Table 4.7: Summary of regression coefficients in predicting job performance by *Model 2*. This reports top 10 coefficients ranked on variable importance [266].

IRB		OCB	
Variable	**Coefficient**	**Variable**	**Coefficient**
Freq. of Conflict Situations	0.59	Adaptability/Flexibility	-49.92
Service Orientation	6.31	Work Schedules	1.45
Recognition	24.10	Face to Face Discussions	0.36
Independence	-9.93	Importance of Being Exact	-0.46
Responsibility for outcomes	0.89	Coaching Others	-37.43
Working Conditions	-8.58	Instructing	-36.56
Freq. of Decision Making	-10.80	Working with Work Group	-0.003
Enterprising	0.96	Conventional	-167.92
Monitoring Resources	0.80	Support	-72.41
Initiative	-9.20	Maintain Relationships	75.35

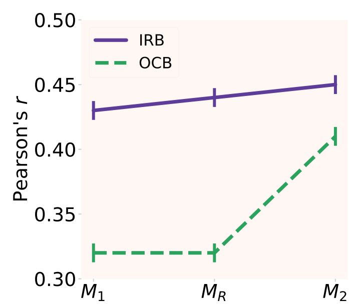


Figure 4.5: Pearson's r of Models predicting individual job performance (M_1: Model w/o *OC*, M_R: Model w/ Org. Sector wise Rating, M_2: Model w/ *OC* via Language)

ratings, I examine if features based on linguistic aspects of the content offer anything more than raw scores. I build a third model where I only replace *OC* in Model Equation *Model 2* with mean aggregated rating per sector. The Ridge model performs the best in both the job performance measure predictions. For IRB, this model shows an adjusted R^2=0.24, Pearson's r=0.43, and SMAPE=3.65.

For OCB, this model shows Adj. R^2=0.14, Pearson's r=0.32, and SMAPE=6.95 — this model performs only as good as *Model 1* (ref: Figure 4.5). So, Glassdoor content when quantified in the lexico-semantic space bears greater explanatory power compared to a single numeric rating. This adds credence to our approach of operationalizing o *OC* as a multi-dimensional construct [445] instead of relying on a single value.

4.2.4 Discussion

This study presents a novel methodology to quantitatively model *OC* through crowd-contributed employee experiences at workplaces. This study reveals that crowd-contributed workplace experiences on anonymized review platforms such as Glassdoor explains the lexico-semantics of *OC*. That is, the dimensions of organizational culture explain individuals' perceived outlook and experiences about the company. Further, situating our findings with the Social Ecological Model [102], this study reinforces concepts in organizational behavior research by validating that this model can significantly explain individual performance at the workplace. This study bears implication in understanding workplace experiences, and in designing empirically guided data-driven technologies to help improve organizational functioning.

Theoretical and Methodological Implications

Beyond Surveys/Ratings. By leveraging crowd-contributed experiences of workplaces shared in an unprompted way online, this study mitigates the limitations of traditional surveys in assessing *OC* [133, 135, 240, 304, 495]. While traditional surveys that summarize information into a singular score have its benefits, this compression of information loses the nuance of the multidimensional nature of *OC* [489]. This challenge is tackled by conceptualizing *OC* on the basis of 41 dimensions and their lexico-semantic space. Another challenge of traditional surveys at the workplace is their vulnerability to a number of response biases [33, 45, 631]. In an organizational setup, a participant's privacy insecurities of getting exposed to management are amplified [236, 321]. This leads to both social desirability and non-response bias. Moreover, many individuals with counter-views and unpopular opinion do not end up participating in such studies to begin with, unevenly skewing the data. Alternatively, this study uses data from Glassdoor where content is public, anonymous, and not actively solicited [244]. Although prior experience and personality can affect public disclosure online, here it is primarily driven by altruism, knowledge, and self-efficacy [120, 362, 369]. Surveys are also limited by when and how frequently they are conducted, whereas *OC* gradually evolves over time [16]. Unlike surveys, this method can used in splices of time to empirically trace the cultural evolution of organizations [600].

Organizational Culture as a Linguistic Construct. This study contributes a word-vector based lexico-semantic similarity approach to model *OC*, furthering earlier approaches of modeling *OC* [10, 215, 439, 464]. The lexico-semantics of language capture the underlying cultural setup of an organization [39, 250, 276]. Although review platforms have traditionally been considered as a mechanism to rank, compare, or recommend across entities like companies, our work provides evidence that anonymized (but well-moderated) platforms such as Glassdoor can be leveraged as a reflection of offline and/or situated communities [47]), and their norms and practices.

Practical and Design Implications

Interest in the topic of human resource management is still nascent in the HCI, and cross-disciplinary literature pertaining to workplaces and technology provides several use cases urging the attention of designers [154, 574]. Along these lines, this study presents opportunities to design for personnel management and organizational decision making. Building on the notion that *OC* is associated with job attractiveness [93], I discuss employee- and employer- facing implications below.

"How is It Like Working in Company X?" Modeling *OC* can render a normative "signature" of an organization, which can feature on public online platforms. This can help both job-seekers and existing employees to reflect on the assumptions and expectations of a work environment [159, 574]. Additionally, enterprise-based social-networks and collaborative knowledge bases or "wikis" are already used by employees to learn about and engage in their organization's culture [11, 105, 615]. Integrating language-based models of *OC* on these platforms can help teams understand the work environment and beliefs and attitudes within an organization.

"How Healthy is Our Culture?" From an employer's perspective, an actionable representation of *OC*, delivered through privacy-preserving, employee-aware technologies and interfaces, can provide a concrete sense of both individual and collective performance. *OC* arguably has a strong influence on employee behavior [39, 445]. This study can help companies to unpack the atmosphere developing within the workplace through questions like: "Does our culture support work-life balance? Does our organization enhance employee creativity? Or is it concentrated only on productivity? Do we celebrate, incentivize, reward, recognize individual efforts well enough? Do we have enough collaboration? Do employees enjoy doing that?" Importantly, with an ability to gauge *OC*, companies can inspect how well leadership structures model behaviors that embody the company culture, how important events (e.g., IPOs, product releases), may impact the culture, and what steps might address issues of unhealthy culture.

Ethical Implications and Considerations

Meaningfulness of Glassdoor Data: Bias and Abuse. Although this method is agnostic to the nature of the platform, it is undeniable that its credibility and consequences in a practical deployment hinge on the characteristics of the platform. Glassdoor claims to be equitable in its moderation and presentation of different reviews irrespective of ratings [244]. Even though they champion free speech, they avoid platform abuse and illegitimate skewing / polarization of reviews, they establish strict user limitations. However, guidelines can be breached and even updated. Even with clear guidelines, users can develop behaviors that are "within the rules" but may deter the overall meaningfulness of the data. Despite content balancing policies like "give to get" [242], review sites like Glassdoor can face *retaliatory utilization* — where dissatisfied employees are more likely to post [318]. A similar problem is intentional tarnishing of employers by trolls. Since our work shows the applicability of data from online platforms to understand offline organizational constructs, it should motivate stronger policies to avoid misbehavior and presence of "bad actors".

Reputation Building and Divergent Views. This study makes us being critical regarding whether small organizations are as empowered as the ones with large volume of users and history. Companies with more employee reviews will be robust to diverse opinions and therefore may find it easy to build and maintain their reputation on public platforms. This study also recognizes that smaller companies, especially those in early stages, may find it challenging to build a reputation when it can be easily misconstrued with a few extreme reviews. In fact, given representation of *OC*, potential employees could leverage inappropriate portrayals of smaller companies' cultures as an extortion tactic to negotiate pay and benefits [611, 672]. This study encourages online platforms to consider new ways to protect organizational profiles.

Manipulative Intent to Alter Cultural Perception. Formulating *OC* ignores the nuances of user behavior on sites like Glassdoor [183, 582, 682]. Admittedly, these vulnerabilities can be exploited to harm a company's reputation, and alternatively organizations may game the system to boost attractiveness. Crowd-contributed platforms in other spheres like service and product feedback are rife with problems of "review fraud", where artificial reviews alter public perception of products [384, 433]. Similarly, this study can be abused to selectively manipulate information and jeopardize employee agency, such as by discouraging posts that with less desirable cultural attributes, and consequentially harm, or even socially alienate the employees who identify with those attributes.

4.3 Modeling Role Ambiguity with Online Professional Porfolios

The complexities related to an individual's job role, or the *expectations applied to an individual within and beyond an organization's boundaries* can impact their job satisfaction [639]. In fact, any sort of discrepancy between *what* an employer expects and *what* an employee does at the workplace can impact wellbeing and performance, as employees can find themselves pulled in various directions as they try to respond to the many statuses they hold. According to the “Role Theory”, role conflict, role ambiguity, and role overload are three aspects of job role that contribute to workplace stress, or the stress that arises if the demands of an individual’s roles and responsibilities exceed their capacity and capability to cope [336, 471]. Among the role constructs, role ambiguity has been considered to be the most significant one, and it is also the focus of the current study [336].

Role ambiguity is broadly considered to include uncertainties about role definition, expectations, responsibilities, tasks, and behaviors involved in one or more facets of the task environment [320, 336, 558]. Role ambiguity has both objective and subjective components — Objective role ambiguity refers to external conditions in the individual’s workplace environment, whereas subjective role ambiguity relates to the amount of ambiguity that the individual perceives in their workplace owing to the information gap that they face [336]. Further, role ambiguity leads to consequences related to dissatisfaction, distrust, lack of loyalty, turnover, absenteeism, low performance, anxiety-stress, and increased heart rate [639]. There is sufficient evidence demonstrating how role ambiguity negatively affects one’s organizational life in terms of their physiological, behavioral, psychological, and performance related measures [335, 558].

Traditionally, role ambiguity is measured using survey instruments that record employee responses to their perceived clarity of assigned tasks, expectations on the job, expectations of peers, and if these peers explicitly mention their expectations from the focal employee [509]. In particular, these methods not only suffer from subjective biases [583], but also are only able to capture the “perceived” component of role ambiguity. Individuals may or may not be aware that they are working on things beyond their job requirements, such as when there is an information gap, or if they are investing their effort to gain knowledge and experience [230, 345]. Thus, it is unclear how useful these measures are [508], and researchers have argued that the lack of an instrument capable of measuring objective and perceived facets of ambiguity may have impeded both theory development and application of research results [76].

Further, with the development and adoption of technology in several sectors of the workplace, the landscape of work is evolving at an unprecedented speed. This also demands continuous skill development [110, 327]; a recent study by McKinsey Global Institute predicts enormous workforce transitions in the years ahead, estimating by 2030, as many as 375M workers globally will likely need to transition to new occupational categories and learn new skills [395]. However, there is no defined approach to proactively gauge individuals' fit with their assigned roles, no guidance for interventions to help them overcome role ambiguity. An organization that can proactively deal with role ambiguity will benefit from employees with increased satisfaction, wellbeing, and productivity in general.

This study contributes to the above research gap and advances the theory by introducing a novel way of measuring role ambiguity. To the best of our knowledge, our study is the first to empirically and objectively measure role ambiguity via LinkedIn, a professional social networking platform where career profiles are publicly shared by employees with self-descriptions of their job titles and role descriptions. Juxtaposing traditional surveys with modern sensor derived measures of wellbeing, I combine methods adopted from natural language analysis and statistical modeling to examine the relationship of LinkedIn based role ambiguity (`LibRA`) with the wellbeing and job performance of individuals — the two important facets corresponding to one's job satisfaction [558].

Aim 1. To measure role ambiguity using unobtrusively obtained LinkedIn data.

Aim 2. To examine the relationship of Linkedin based Role Ambiguity (`LibRA`) with individual wellbeing and job performance.

Aim 3. To investigate what factors may contribute to one's `LibRA`, relating to their intrinsic traits, LinkedIn's platform-specific characteristics, and preferences and goals of use of professional social networking service.

Towards Aim 1, I model LinkedIn based Role Ambiguity (`LibRA`) as a lexico-semantic difference between the job description of an individual's role as self-portrayed on their Linkedin profile and what is posted by the company for that particular role. I employ natural language analysis to obtain word embeddings of job descriptions along eight facets of job role, namely *abilities, interests, knowledge, skills, work activities, work context, work styles, and work values* [610]. Towards Aim 2, I test for theory-driven hypotheses that examine the relationship of `LibRA` with an individual's 1) wellbeing related measures, namely their heart rate, arousal, sleep, and work-hours, and 2) job performance related measures, namely their individual task performance, in-role behavior, and organizational citizenship behavior. For Aim 3, I reflect back to investigate factors contributing to `LibRA`. I contextualize one's self-presentation behavior on LinkedIn to draw insights into the unobservable and unaccounted factors, such as an individual's mindset, job-related motivation, and platform-related nuances.

4.3.1 Data: LinkedIn

LinkedIn Data

Out of the 757 participants in the study, 529 provided their LinkedIn data. Our work accounts for those with self-described portfolios and their passively sensed and self-reported wellbeing and job performance data. Therefore, I filter out "blinded" participants and those without any self-description in their LinkedIn profile, particularly in their profile and job summary, leading us to a LinkedIn dataset of 257 individuals — all the ensuing analyses in this study is limited to these 257 individuals' data. Corresponding to every participant, the Tesserae project obtained their self-presented profile and job summary which includes current and previous jobs. Figure 4.6(c) shows the top job titles in our dataset, and Figure 4.6(d&e) shows two word-trees of profile summaries on two top representative keywords ("professional" and "skill") in our dataset. These word-trees hint how individuals self-present their job summaries on their LinkedIn profiles; for example, within skills, we find occurrences of both tangible/technical skills (eg. *sap, technology, sales, microsoft office*) and intangible/people skills (eg. *leadership, communication, analytical, interpersonal*).

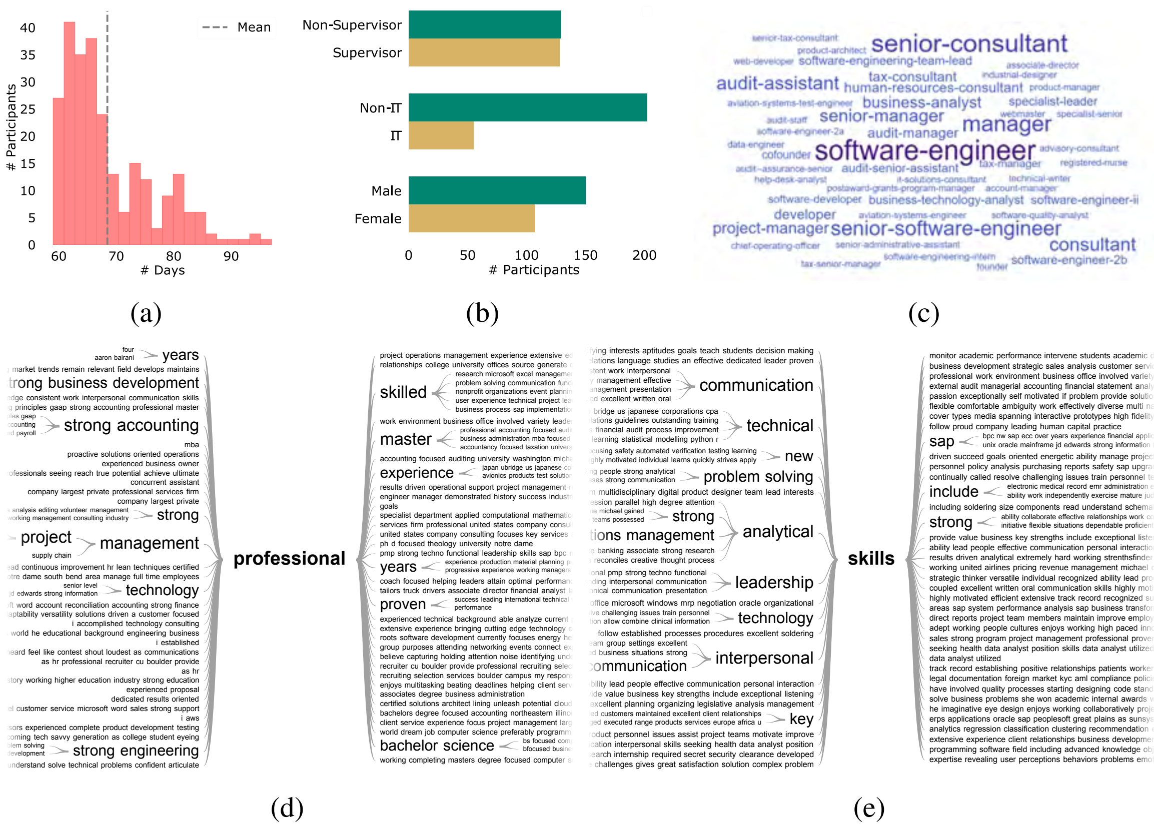


Figure 4.6: Distribution in the dataset on (a) study period per participant, (b) demographic and job-role based characteristics, (c) word-cloud on the job roles on LinkedIn data, (d&e) Word tree visualizations on two top-occurring keywords (professional and skills) in the LinkedIn profile descriptions: These visualizations show content in the form of co-occurrences of keywords in the dataset. The font size of keywords are proportional to their occurrence along with surrounding co-occurring keywords. For example, management profesional, technology professional, professional keywords, analytical skills, technological skills, etc.

Self-Reported Data

Figure 4.6b plots the demographic distribution of the participants who provided us LinkedIn data. The 257 participants with *complete* LinkedIn data consist of 150 males and 107 females. The average age of the participants is 35.2 years (stdev. = 9.5). These participants belong to 60 unique companies, and among these, the top three companies, 103 belong to C_1 (a large-scale multinational firm), 54 belong to C_2 (a mid-size product-centric firm), and 17 belong to C_3 (a research organization). In terms of job role, the data contains 128 supervisors and 139 non-supervisors. In job sector, 202 participants belong to Non-IT sector, and 55 participants belong to the IT sector. In terms of tenure, while a majority of the individuals (53) have been at their current organization for over eight years, 113 individuals have been at their current organization between three to eight years, and 101 individuals have been at their current organization for less than three years. For education, most participants have a college (52%) or master's degree (35%). In income, the participants are more evenly distributed, with the majority (64%) of the participants similarly distributed in the 50K-75K, 75K-100K, and more than 150K USD income brackets.

The Tesserae project obtained the participants' big-five **personality traits** assessed by the Big Five Inventory (BFI-2) scale [589, 614], and **executive function** assessed by the Shipley scale [104, 559, 578]. For personality traits, the dataset shows a mean openness of 3.86 (std.=0.54), mean conscientiousness of 3.87 (std.=0.66), mean extraversion of 3.41 (std.=0.67), mean agreeableness of 3.85 (std.=0.55), and mean neuroticism of 2.51 (std.=0.77). For executive function, the dataset shows a mean fluid intelligence of 33.38 (std.=4.18), and mean crystallized intelligence of 16.91 (std.=2.79).

4.3.2 Aim 1: Measuring Role Ambiguity from LinkedIn (LibRA)

Why LinkedIn? LinkedIn is a professional social networking platform (launched in 2003) that allows individuals to create and publish their professional profiles and describe and their portfolios. Although LinkedIn is biased towards individuals' positive self-presentation and

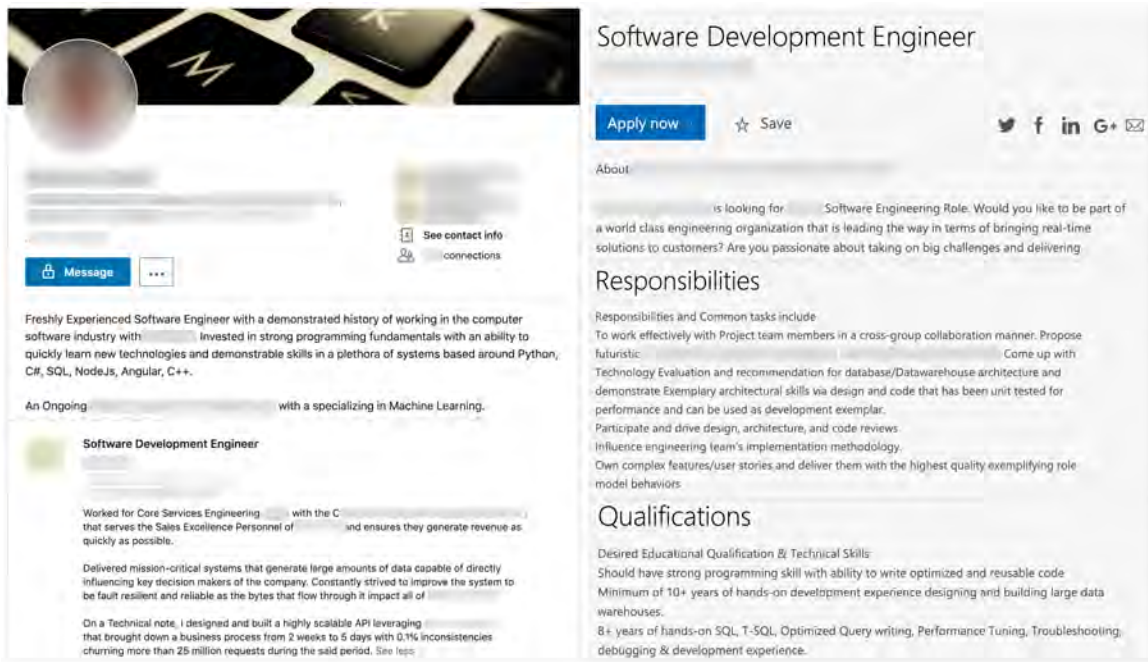


Figure 4.7: For the same role (Software Development Engineer): (left) Role summary of an individual as described on LinkedIn, (right) Job description as posted on the company webpage

self-promotion, the non-anonymity and public-facing nature of the platform also influences individuals to be less deceptive and more accountable in their profiles [274]. In line with Goffman's theory of self-presentation, LinkedIn provides an ideal platform for individuals to present their "professional" selves to the online audience [249, 633]. Because LinkedIn is a non-anonymous platform, where individual identity (at least the name) is disclosed, it somewhat helps promote trust and accountability on the platform [183]. Therefore, it suits the choice of our dataset where we seek to obtain self-presented portfolios of employees on their roles and responsibilities at organizations.

Libra: LinkedIn based Role Ambiguity

Defining LibRA. Drawing upon the theoretical definition of role ambiguity, we operationalize LinkedIn based Role Ambiguity (`LibRA`) as the *quantified differences in the self-explained roles and responsibilities of the individual against that posted by the company* *for the same role in the organization*. For this, we obtain the self-explained job summary from an individual's LinkedIn profile. Then, for each role, we obtain the company described job description by manually conducting search engine queries of the specific role and the company. These job descriptions are typically posted on job posting websites, such as *Glassdoor, LinkedIn, Indeed*, and the *Google job search portal* — where the Google job search portal collates both exact and nearest matching job descriptions from multiple websites, including company's own website, LinkedIn, Glassdoor, Indeed, etc, and sorts them in relevance to the search query. Figure 4.7 shows an example LinkedIn role description and company-published role description for the same role of Software Development Engineer at the same location of the company. Two members of the research team independently obtained the nearest matching job description per role and per company — there were very few ($<$20) instances when the two coauthors obtained different job descriptions, and when they did, the descriptions were very similar in the two websites, and the more relevant one was chosen.

Assessing LibRA. Towards computing `LibRA`, this study represents the above descriptions of self-reported LinkedIn job descriptions and the company described job descriptions in a multi-dimensional space of job aspects, for which O*NET is used. O*NET[2] is an online database and job ontology that contains a comprehensive list of jobs and their descriptions, elaborating on eight notable aspects of job role — *abilities, interests, knowledge, skills, work activities, work context, work styles, work values* (see Table 4.8 for brief descriptions). These aspects are grounded in literature and have been used in prior work to study employee behavior [610]. For every individual's role, I obtain their closest matching O*NET roles. I adopt a semi-automatic approach of edit-distance based match, followed by manual evaluation and curation by the research team, which is familiar with and are users of LinkedIn. For example, the closest match of *Software Development Engineer* is *Software Developers*.

[2]O*Net (*onetonline.org*) is developed under the sponsorship of the U.S. Department of Labor/Employment and Training Administration (USDOL/ETA).

Table 4.8: Job aspect types with their descriptions as obtained from O*Net.

Job Aspect	Description
Abilities	Enduring attributes of the individual that influence performance.
Interests	Preferences for work environments and outcomes.
Knowledge	Organized sets of principles and facts applying in general domains.
Skills	Developed capacities that facilitate learning or the more rapid acquisition of knowledge.
Work Activities	General types of job behaviors occurring on multiples jobs.
Work Context	Physical and social factors that influence the nature of work.
Work Styles	Personal characteristics that can affect how well someone performs a job.
Work Values	Global aspects of work that are independent to a person's satisfaction.

Figure 4.8: Aspect-wise LibRA for a random set of 50 participants in two companies C_1 (above) and C_2 (below). These visualizations are an example comparison of LibRA within- and across- company employees

Then, drawing on natural language analysis methods, I use word-embeddings, particularly pre-trained GloVe vectors [476, 540] to project the role descriptions of individuals and companies in a 50-dimensional word-vector space, so as to obtain rich lexico-semantic context surrounding the hand-curated job descriptors above [535]. I use cosine similarities to obtain two vector projections in the eight-dimensional job aspect space per individual i—1) one that is obtained from their LinkedIn summary (v_1^i) and 2) one that is obtained from the same role's company description (v_2^i). Then, the overall LibRA is measured as the euclidean distance between v_1^i and v_2^i. To obtain the aspect-wise LibRA of an individual as the absolute difference per dimension of v_1^i and v_2^i. For instance, Figure 4.8 show heatmaps of multi-dimensional role ambiguity of randomly selected 50 individuals from two companies, C_1 and C_2 in our dataset. We find that some individuals show high or low role ambiguity across the aspects, but most individuals show high role ambiguity in one or more dimensions. While exploring the differences across multi-dimensional role ambiguity constructs remain a future research goal, such multi-dimensional role ambiguity [581] can benefit various stakeholders (employers or employees) through guided intervention to minimize their role ambiguity. This kind of interface is additionally inspired from prior HCI work aimed at facilitating employee satisfaction [159, 573].

Evaluating the Validity of LibRA Against Gold Standard

After defining and proposing a method to measure `LibRA` using LinkedIn data of individuals, I examine the validity of the measure. That is, I examine if the `LibRA` measure gets at least close to what the Role theory identifies as "role ambiguity". For this, drawing on modern validity theory [145], I compare the `LibRA` of the individuals against a gold standard validated survey on measuring role ambiguity. The Michigan Assessment of Organization survey instrument measures an individual's role ambiguity, role conflict, and role overload [435]. Corresponding to role ambiguity, the scale asks the participants to rate the four statements, "Most of the times I know what I have to do on my job", "On my job I know exactly what is expected of me", "I can usually predict what others will expect of me on my job", and "Most of the time, people make it clear what others expect of me", a 7-point Likert scale ranging from "Strongly Agree" to "Strongly Disagree".

I randomly sample a subset of 77 participants from our entire participant pool to answer the Michigan Assessment of Organization survey [435]. Correlating the survey-based role ambiguity with `LibRA`, I find Spearman's[3] correlation coefficient to be 0.22 ($p < 0.05$).

Consequently, a statistically significant correlation does imply criterion validity, and hints at construct validity in our claim that `LibRA` does contain information that is also captured by gold standard, validated survey instruments on role ambiguity. However, I also acknowledge that the magnitude of correlation is moderate, which could be attributed to the differences in the measures (one is "perceived", and other is objectively measured).

[3]Because the survey instrument on role ambiguity and our measure of `LibRA` measure role ambiguity in different scale and order, it makes sense to correlate the ranked (or relative) values rather than the raw values

Table 4.9: Summary of covariates used in the regression models.

Covariates	Value Type	Values / Distribution
Demographic Characteristics		
Gender	Categorical	Male | Female
Age	Continuous	Range (22:63), Mean = 35.24, Std. = 9.46
Education Level	Ordinal	4 values [College, Grad., Master's, Doctoral]
Job-Related Characteristics		
Income	Ordinal	7 values [<$25K, $25-50K, ... , >150K]
Tenure	Ordinal	10 values [<1 Y, 1Y, 2Y, ... 8Y, >8Y]
Supervisory Role	Boolean	Supervisor | Non-Supervisor
Job Type	Boolean	IT | Non-IT
Executive Function (Shipley scale)		
Fluid (Abstraction)	Continuous	Range (5:23), Mean = 16.91, Std. = 2.78
Crystallized (Vocabulary)	Continuous	Range (0.0:40.0), Mean = 33.38, Std. = 4.18
Personality Trait (BFI scale)		
Openness	Continuous	Range (1.7:5.0), Mean = 3.86, Std. = 0.54
Conscientiousness	Continuous	Range (1.7:5.0), Mean = 3.87, Std. = 0.66
Extraversion	Continuous	Range (1.7:5.0), Mean = 3.41, Std. = 0.67
Agreeableness	Continuous	Range (2.1:5.0), Mean = 3.85, Std. = 0.56
Neuroticism	Continuous	Range (1.1:4.7), Mean = 2.51, Std. = 0.77

4.3.3 Aim 2: Examining Relationship of LibRA with Wellbeing and Performance

Theoretical Underpinnings and Hypotheses

Role Ambiguity and Wellbeing. While there is no single conceptualization of wellbeing, the broad categories that wellbeing encompasses are physiological, psychological and behavioral aspects [335, 558]. Physiological indicators include factors such as blood pressure, heart conditions, and general physical health. Psychological indicators include affect, frustration, anxiety, stress, and arousal. Behavioral aspects include those that an employee has a choice to make, like the time spent at work, the time taken for breaks during work, mobility to another employment (turnover), hours of sleep, etc.

Within the scope of our dataset, I study the relationship of `LibRA` with one's physiological measures (heart rate and sleep [99, 113]), psychological measures (stressful arousal [605]), and behavior at the workplace (time spent at desk and time spent at workplace [679]). Specifically, I test for the following hypotheses in the relationship of `LibRA` with wellbeing attributes.

H_1. Greater role ambiguity is associated with increased heart rate.

H_2. Greater role ambiguity is associated with increased arousal.

H_3. Greater role ambiguity is associated with decreased sleep.

H_4. Greater role ambiguity is associated with reduced work-hours.

Role Ambiguity and Job Performance. Role ambiguity consists of the uncertainty regarding tasks that an employee needs to perform as part of their job role in the company. An employee with greater clarity will be able to better perform the required tasks. One plausible mechanism that can explain this higher performance is the intrinsic motivation of an employee [387, 509]. Lower role ambiguity or higher role clarity makes it easier to meet the expectations, the employee more motivated and such intrinsically motivated employees perform better and more efficient [229, 231]. Employees with higher job satisfaction are intrinsically motivated and strive harder at work which contributes to their performance. Thereby, the exposure to role stressors (such as role ambiguity) affects an individual's capacity to control their work environment, which in turn adversely affects their ability to function effectively [365, 410].

Within the scope of our dataset, I study the relationship of `LibRA` with two dimensions of job performance [519, 643, 663] — 1) task performance and 2) organizational citizenship behavior. Prior literature in organizational behavior dominantly uses these subjective measures and I rely on the extant literature for the validity of these measures [643]. I test the following hypotheses for `LibRA` with respect to job performance.

H_5. Greater role ambiguity is associated with decreased task performance.

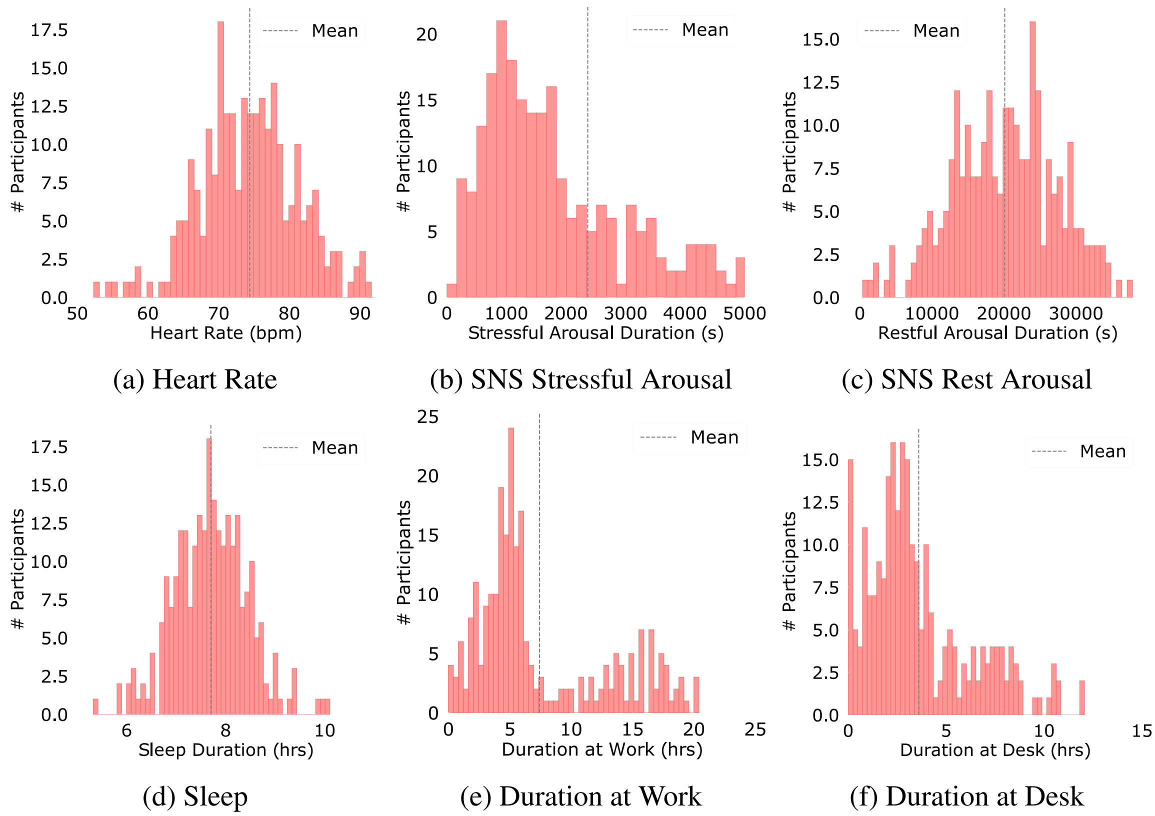
H_6. Greater role ambiguity is associated with decreased organizational citizenship.

Testing Hypotheses and Convergent Validity of `LibRA`

To establish the convergent validity of `LibRA`, I adopt a theory-driven approach to outline hypotheses on the relationship of `LibRA` with job performance and wellbeing. For this, I study the relationship (and association) of `LibRA` with the passively sensed wellbeing measures, and the validated survey-based job proficiency measures. This study uses linear regression models, which are known to provide easily interpretable associations in cases of conditionally monotone relationships with the outcome variable [158]. For every wellbeing or performance measure $\mathcal{M}$, I build linear regression models with $\mathcal{M}$ as the dependent variable, and `LibRA` as an independent variable, controlled for demographic, personality, and executive function measures per individual (see Equation Equation 4.1). Our choice and inclusion of these covariates are motivated from prior literature [42, 72, 671]. Table 4.9 summarizes these covariates in their kind, and the values attained. For all the regression models, I use variance inflation factor (VIF) to eliminate multicollinearity of covariates (if any) [446]. For the ease of comparing the relative importance of the predictive variables in the regression models, I standardize the variables such that they have a mean of zero and standard deviation of one.

$$\begin{aligned}\mathcal{M} \sim\ & gender + age + education\_level + income + supervisory\_role + tenure + job\_type \\ & + executive\_function + personality\_trait + LibRA\end{aligned} \tag{4.1}$$

Testing H_1: Greater role ambiguity is associated with greater heart rate High heart rate is associated with an increase in stress [56, 296]. Caplan and Jones found that greater role ambiguity is associated with increased heart rate, which is identified as a major predictor of coronary heart rate [56, 99]. I obtain the heart rate measures of the participants through the wearable sensor (see section 4.1, Figure 4.9 (a)). I fit a linear regression model with the average heart rate (HR) in the study period per individual. Given that exercise and physical activity has an association with heart rate [254], in addition to the covariates listed in Table 4.9, I control for the physical activity per participant. The regression model reveals a positive standardized coefficient (0.10) with statistical significance for `LibRA` (Table 4.10, Figure 4.10 (a)). This observation supports our Hypothesis H_1.



(a) Heart Rate (b) SNS Stressful Arousal (c) SNS Rest Arousal

(d) Sleep (e) Duration at Work (f) Duration at Desk

Figure 4.9: Distribution of wellbeing measures as inferred via passive sensors.

Testing H_2 Greater role ambiguity is associated with greater stressful arousal Arousal is a physiological response that is related to one's heart rate variability, and is associated with stress, fatigue, and anxiety [176, 296]. These wellbeing measures are known to exacerbate in the presence of role ambiguity [4, 99]. In our project, the wearable sensor allows us to obtain participant arousal, particularly their Sympathetic Nervous System (SNS) arousal measures in a continuous fashion. In particular, for every individual, it scores the arousal level from restful to stressful on a scale of 1-100 at every three-minute granularity (Figure 4.9 (b&c)). Here, the restful duration is when an individual relaxes or recovers from stress [234]. I build two separate regression models, one with median duration of high stressful arousal (75-100), and one with median duration of restful arousal (1-25) per individual. We find that `LibRA` shows a positive standardized coefficient (0.42) in the former model, and a negative standardized coefficient (-0.22) in the latter model (Table 4.10, Figure 4.10 (b&c)). This suggests that individuals with high `LibRA` are more likely to show higher stressful arousal, and lower restful arousal. Therefore, our observations support H_2.

Testing H$_3$: Greater role ambiguity is associated with decreased sleep Sleep is an important attribute in an individual's wellbeing, and it reduces the negative impact of stress as well as improving overall health [67]. Given that stress reduces sleep, and sleep reduces stress, a stressed person is likely to sleep less [636]. If role ambiguity is stressful, this study hypothesizes that high role ambiguity will correspond with reduced sleep duration. The wearable sensor allows us to obtain participant sleep durations (see Figure 4.9 (d)). I build a linear regression model with median duration of sleep per individual. We find that `LibRA` shows a negative standardized coefficient (-0.16) with statistical significance (Table 4.10, Figure 4.10 (d)). Therefore, H_3 is supported in our dataset.

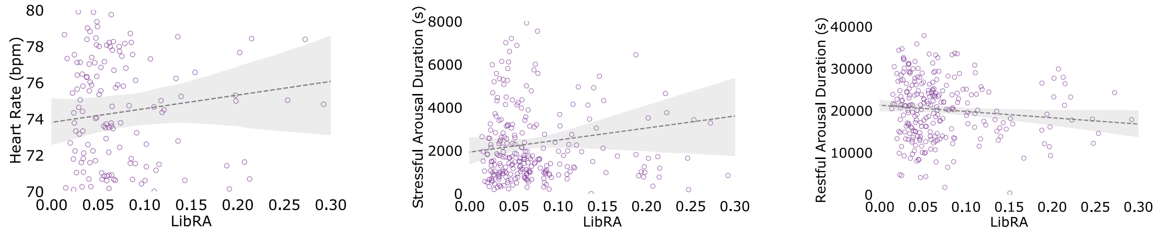
Testing H$_4$: Greater role ambiguity is associated with decreased work hours Role ambiguity is known to affect an individual's workplace behavior [471]. The bluetooth beacons sense if a participant is at work, at home, or commute, and within work; it additionally captures the duration the participant is at- and away from- desk. I build two regression models, one with the duration at work, and one with the duration at desk, when at work (this model additionally controlled for duration at work). Here, we find that both of these dependent variables show heavy-tailed distributions (see Figure 4.9 (e&f). For both of these distributions, Chi-squared tests could not reject the null hypotheses that they were significantly different from a Poisson distribution ($p > 0.05$). Therefore, instead of using purely linear regression models, I build negative binomial regression models [300], ones that essentially regress the logarithm of the dependent variables with the independent variables [300]. Negative binomial regression is preferred over poisson regression because we find the presence of over-dispersion in the distribution of both duration at work and duration at desk (Figure 4.9 (e&f)) [151]). `LibRA` shows a negative standardized coefficient in both the models (-0.41 for duration at work, and -0.12 for duration at desk, Table 4.10, Figure 4.10 (e&f)). This

Table 4.10: Summary of standardized coefficients of regression models of wellbeing.

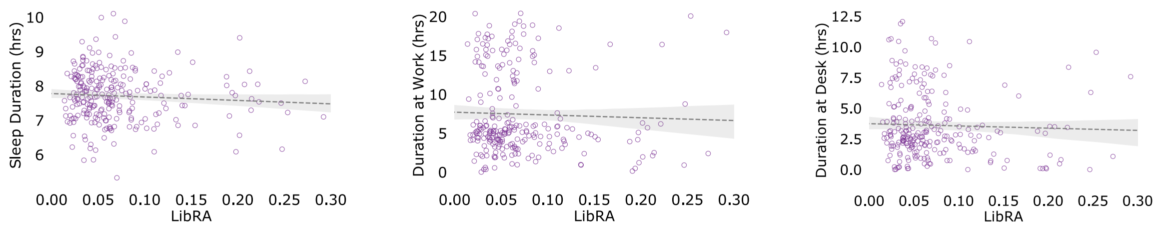
Covariates	Std. Coeff.	Covariates	Std. Coeff.	Covariates	Std. Coeff.
$\mathbf{H}_1$ (Heart Rate)		$\mathbf{H}_2$ (Arousal)			
$\mathcal{M}$ = Heart Rate, $R^2 = 0.16$*		$\mathcal{M}$ = Stressful Duration, $R^2 = 0.65$*		$\mathcal{M}$ = Restful Duration, $R^2 = 0.47$**	
Exercise Duration	0.53**	Age	0.69**	Job: Non-IT	0.31*
Shipley: Abs.	-0.81*	Edu: Grad. School	-0.24*	**LibRA**	-0.22***
Agreeableness	0.91*	Tenure: 4	-1.59*		
Conscientiousness	-0.78*	**LibRA**	0.42***		
LibRA	0.10*				
$\mathbf{H}_3$ (Sleep)		$\mathbf{H}_4$ (Work-Hours)			
$\mathcal{M}$ = Sleep Duration, $R^2 = 0.19$***		$\mathcal{M}$ = Duration at Work		$\mathcal{M}$ = Duration at Desk	
Income: $50K-75K	0.21*	Edu.: College	0.23***	Duration at Work	0.18*
Agreeableness	-0.14*	Edu.: Grad. School	0.21***	Edu: College	-0.09
Tenure: 7 Yrs.	-1.74*	Income: $50K-75K	0.14***	Edu: Grad.	-0.04
Job: Non-IT	0.15**	Income: $100K-125K	-0.18***	Edu: Master's	0.04
LibRA	-0.16***	Shipley: Abs.	0.01***	Income: $100K-125K	0.09*
		Extraversion	0.09***	Income: $125K-150K	0.08*
		Conscientiousness	0.05***	Tenure: $<$1 Yr.	-0.18***
		Neuroticism	0.12**	Tenure: 2 Yrs.	0.18***
		Tenure: 6 Yrs.	-0.16***	Tenure: 3 Yrs.	0.26***
		Tenure: 7 Yrs.	-0.15***	Tenure: 4 Yrs.	0.09***
		Tenure: 8 Yrs.	-0.31***	Tenure: 8 Yrs.	0.15***
		Job: Non-IT	0.20***	Job: Non-IT	-0.03*
		LibRA	-0.41***	**LibRA**	-0.12**

suggests that individuals with high `LibRA` are not only less likely to spend time at work, but also less likely to spend time at desk when at work. These observations support H_4.

Testing H_5: Greater role ambiguity is associated with lower task performance Two survey scales of In-Role Behavior (IRB) and Individual Task Performance (ITP) three times a week were administered, to periodically obtain the self-assessed task performance of the participants (see Section section 4.1, Figure 4.11 (a&b)). For both these measures, I build two linear regression models each — one that uses an aggregated (median) value of task performance, and one that uses a change in task performance over the duration of the study. We find that `LibRA` shows a negative association with both *aggregated ITP* (-0.33) and *change in ITP* (-0.20) per individual. Similarly, `LibRA` also shows a negative association with both *aggregated IRB* (-0.29) and *change in IRB* (-0.20) per individual (Table 4.11, Fig.Figure 4.12 (a&b, d&e)). Together, these observations suggest that individuals with higher `LibRA` not only have a greater likelihood of performing badly at work, but also their performance worsens over time. Therefore, our observations support H_5.



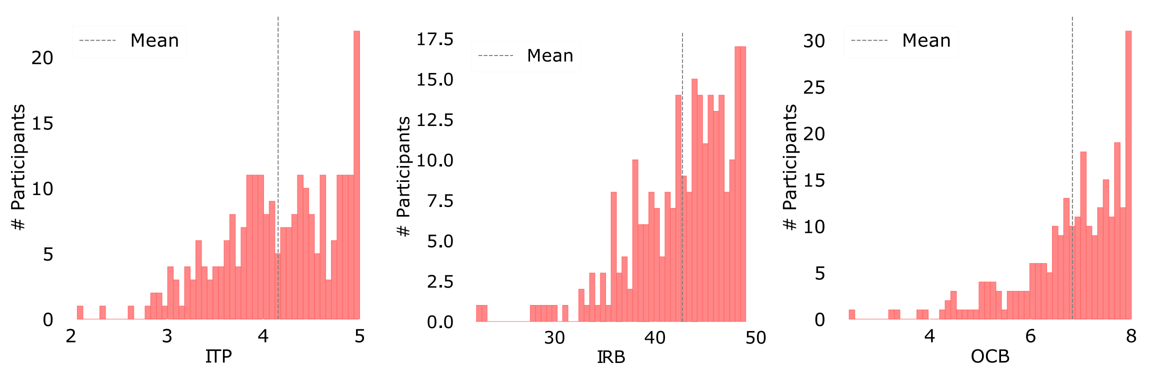
(a) Heart Rate (b) SNS Stressful Arousal (c) SNS Restful Arousal



(d) Sleep (e) Duration at Work (f) Duration at Desk

Figure 4.10: Scatter plots of demonstrating the distribution of wellbeing attributes against `LibRA`. `LibRA` is positively associated with heart rate, stressful arousal, and negatively associated with restful arousal, sleep, duration at work, and duration at desk. In sum, increase in `LibRA` is associated with depleted wellbeing.

Testing H$_6$: Greater role ambiguity is associated with lower organizational citizenship behavior We administered the Organizational Citizenship Behavior (OCB) scale three times a week, to periodically obtain the self-assessed organizational citizenship behavior of the participants (Figure 4.11 (c)). Like the above, I build two linear regression models — one that uses an aggregated (median) value of OCB, and one that uses a change in OCB over the duration of the study. We find that `LibRA` shows a negative association with both



(a) ITP (b) IRB (c) OCB

Figure 4.11: Distribution of Performance measures via job performance surveys.

Table 4.11: Summary of standardized coefficients of regression models of task performance.

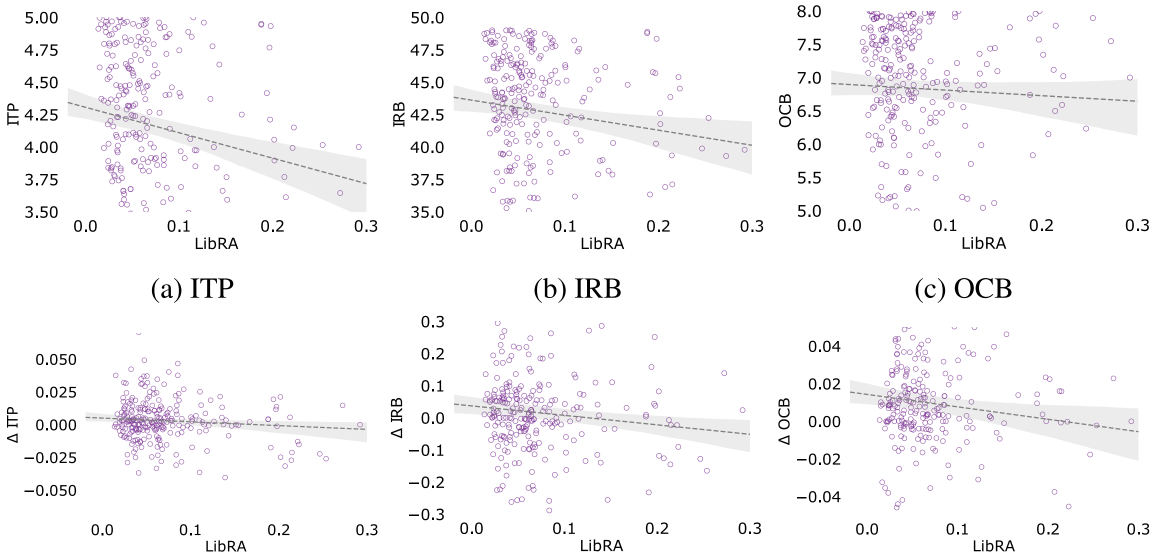
Covariates	Std. Coeff.	Covariates	Std. Coeff.	Covariates	Std. Coeff.
H_5 (Task Performance)				H_6 (Org. Citizenship Behavior)	
$\mathcal{M}$ = ITP, $R^2 = 0.29$***		$\mathcal{M}$ = IRB, $R^2 = 0.29$***		$\mathcal{M}$ = OCB, $R^2 = 0.24$***	
Income: L	-0.38*	Openness	0.13**	Supervisor: Yes	0.24***
Income: Q	0.40**	Consc.	1.13*	Extraversion	0.34***
Openness	1.07*	Tenure: 8	0.17*	Tenure: 6	-0.14*
Consc.	1.30***	**LibRA**	-0.29*	Tenure: 7	-0.20**
Tenure: 6	-0.15*			**LibRA**	-0.10**
LibRA	-0.33***				
$\mathcal{M}$ = Δ ITP, $R^2 = 0.13$*		$\mathcal{M}$ = Δ IRB, $R^2 = 0.17$***		$\mathcal{M}$ = Δ OCB, $R^2 = 0.22$*	
Extraversion	0.69*	Openness	0.91**	Supervisor: Yes	-0.26*
Consc.	-1.37***	Consc.	-0.84*	Agreeableness	-1.80*
LibRA	-0.20*	Tenure: 7	-0.19*	Tenure: 5	0.21*
		Tenure: 8	-0.26**	**LibRA**	-0.25***
		Tenure: 9	-0.18**		
		LibRA	-0.20**		

aggregated OCB and change in OCB per individual (Table 4.11, Figure 4.12 (c&f)). These observations suggest that individuals with higher `LibRA` show a greater likelihood of poorer OCB, which also worsens over time — a tendency associated with being disinclined to be altruistic or help colleagues at workplace. Therefore, our observations support H_6.

4.3.4 Investigating the Factors Affecting LibRA

This final section studies the factors that contribute to the LinkedIn based role ambiguity (`LibRA`) assessment. Specifically, I investigate the extent to which appropriating data shared online (on a professional social networking service, LinkedIn) may bring forth new dimensions to consider while employing `LibRA` for practical use, and what might contribute to observed differences in `LibRA`. To do this, I draw from various literature to situate our observations.

First, I seek to quantitatively study the relationship of `LibRA` with observable and intrinsic attributes of an individual. Using the same covariates as in Table 4.9, I fit one's `LibRA` as the dependent variable in a series of statistical models. Our rationale to study this rests on prior work that found demographic and intrinsic traits affecting self-disclosure behavior on LinkedIn, which may lead to differences in `LibRA` [282, 619, 623, 642]. I build multiple regression models (both linear and non-linear), but find no significance in the relationship ($p > 0.1$) in either the regression fit or the variable coefficients. ANOVA F-test per covariate



(a) ITP (b) IRB (c) OCB

(d) Δ ITP (e) Δ IRB (f) Δ OCB

Figure 4.12: Scatter plots of demonstrating the distribution of job performance measures against `LibRA`. `LibRA` is negatively associated with ITP, IRB, OCB, Δ ITP, Δ IRB, and Δ OCB. In sum, an increase in `LibRA` is associated with both decreased job performance as well as reduced job performance over time.

and `LibRA` reinforced our confidence in this finding that there is no significant relationship in the variability of observable traits influencing `LibRA`. This aligns with previous literature that claims role ambiguity is independent of individual traits, rather than an outcome of a number of factors such as mentor-mentee relationship, working alliance, organizational structure, and organizational communication [361]. Nevertheless, because `LibRA` is inferred from social media data, specifically LinkedIn, numerous mediators can confound the self-presentation behavior of an individual on LinkedIn (even after controlling for their intrinsic demographic and personality traits). This study delves deeper into this consideration based on a qualitative examination of a sample of our dataset as described below.

I intend to compare and study the self-presentation behavior, accounting for the between-individual differences in self-reported and assessed traits of demographic, personality, executive function, and work role-related characteristics. Therefore, with these characteristics as covariates (see Table 4.9), I draw on matching techniques from causal inference [317, 544] to match individuals using Mahalanobis Distance Matching [270].

This study separately matches pairs of individuals who belong to IT roles, and who belong to non-IT roles. Figure 4.13 plots the pair-wise Mahalanobis distances and the absolute differences in their role ambiguities. We focus on those individuals (shaded region in Figure 4.13) who are similar in their individual attributes but show high differences in their `LibRA` — this study samples the top 10th percentile of pairs of individuals in IT and non-IT each.

Next, among the individuals in the above sample, the research team manually looks at the (public) LinkedIn job and profile descriptions. While these individuals are very similar in their personality, demographic traits, and their role in the company (because of matching), in terms of self-presentation behavior on LinkedIn, we find differences in their style of writing (also highlighted in the Figure 4.13 examples). For example, one writes an extremely short description compared to their matched other, who writes a longer description with much more detail. Another example includes only technical-skills or the tasks that they are assigned at work (e.g., *Java, business development*), compared to their matched other, who additionally describes their non-technical and people skills and abilities (e.g., *accomplished, dynamic*). Given the affordances and the uniqueness of LinkedIn as a professional social networking platform, I deduce a few plausible reasons that can potentially influence the virtual self-presentation of the individuals, and in turn, lead to varied inferred role ambiguity. I discuss these factors, which are not disjointed and could be inter-related:

Individuals' Organizational Behavior. Individuals who are looking for newer jobs or endeavors possibly write a more detailed portfolio on their LinkedIn profiles, whereas individuals who are generally "settled" are not as active in providing detailed descriptions [582]. This could also be a *different type* of job than what they are currently involved at altogether as well. In other words, the settled people may have different jobs currently compared to they were hired, e.g., through promotion or lateral moves within the company. An alternative conjecture could also be that, only a few individuals write and "highlight" their *work experience*, rather than describing their responsibilities and tasks at the workplace, for

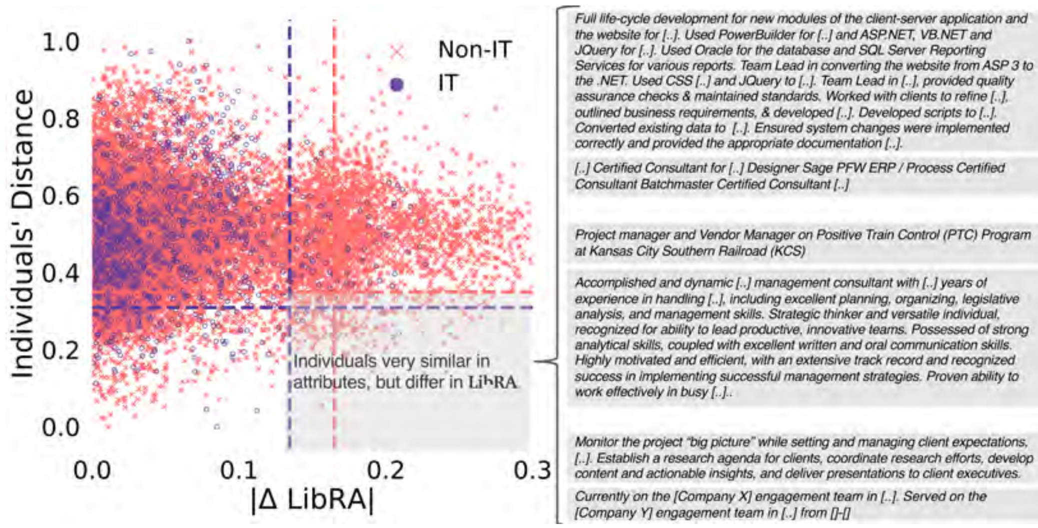


Figure 4.13: Pair-wise differences in individual attributes and corresponding differences in `LibRA`. Example excerpts show differences in LinkedIn descriptions of pairs of individuals with very similar individual attributes (low differences), but large differences in `LibRA`.

example, *"I have 25 years of Health Care Provider experience in revenue cycle selling and managing outsourced health care accounts, receivable solutions [..]"*. We find individuals who describe their role with people skills and proficiencies beyond their tasks, such as *"I can effectively cope with change, shift gears comfortably, and bring a point of view to the leadership"*, and those who describe their attitude towards initiativeness, *"I am always willing to help especially if there is a problem to be solved, and my behavior is a mix of light-heartedness and a drive to put into practice everything I have learned."*. These could be individuals who exhibit proactive behaviors in the organizations [143]: they show anticipatory, change-oriented and self-initiated behavior in situations and tend to act in advance of a future situation, rather than just react later. This may also indicate that although these individuals have high role ambiguity, they show desirable individual characteristics (proactive behavior and leadership traits) in organizations [46, 143]. Exploring these aspects further is of future research interest.

Individual-Intrinsic Factors Prior research has observed that people may self-promote and appear honest and less deceptive on their professional social networking profiles [274, 633]. However, the degree and the way in which they self-present themselves can vary. Given the context of professional choices and career development, we can look at it from the perspective of growth versus fixed mindset [192]. Those with "fixed mindset" believe their abilities are innate, whereas the ones with "growth mindset" believe that abilities can be acquired via investing effort and study. For instance, an individual describes themself as, *"a motivated and hardworking professional looking to improve my skills and abilities."* Although mindset and personality traits are somewhat related, mindset can reshape over time and through interactions [12]. Complementary research directions have also coined "benefit mindset", and "global mindset", "productive mindset", and "defensive mindset", all illustrating a variety of intrinsic behaviors of individuals that contribute to their skill development, proficiency, and self-presentation in organizations [86, 281]. The similar traits likely permeate into online self-promotion practices on LinkedIn.

Job-Related Factors Literature has demonstrated the importance of job titles in organizations [627]. `LibRA` assessments of an individual are derived from the job titles of the individuals. However, if the job titles themselves are ambiguous then that inherently adds ambiguity to the role of the individual. In fact, we find pairs of individuals where one is an "Associate", while the other is a "Specialist" — both of these titles are pretty generic, and do not convey much information to the employees. In contrast, the fact that recently companies are coming up with "cool" job titles (e.g., *ninja*) to gain visibility and distinctiveness can add other complexities to role ambiguity [551]. As Utz and Breuer recently noted that one's career orientation, type of role or organizational sector, influence their behavior or use of LinkedIn— for example some sectors may require more referrals or information than others, thereby implicitly demanding greater activity from the individuals [629, 685]. Additionally, some individuals may be working on confidential projects and they are bound by nondisclosure agreements. Further, the role in a company and size of a company can influence the self-description behavior of individuals [685]. That is, even with common and similar job titles, individuals at large enterprises may not feel the need to describe their role in as much as detail as those at startups and mid-size organizations [402]. Compounding this difference in company size, some companies may encourage the use of LinkedIn among employees to improve the image of the company, or may even render the platform as a mandatory in-company communication tool, thereby influencing the LinkedIn use behavior of their employees [633].

Audience, Privacy, and Platform Factors Finally, the familiarity or the use of LinkedIn as a platform may vary across individuals. Two participants in our sample described what their company does, rather than what their role is, such as, *"[Company] specializes in [..] and works with companies that offer [..] service. [Company] has over 40 years of experience in the industry and operates groups of 10 to 1000 people [..]."* In addition, LinkedIn is a professional social networking platform that also functions as a marketplace for job seekers. Individuals tend to share credible information because they have a conceptualization of an "invisible audience" [57], and since LinkedIn is a public space, they do not want to appear as dishonest [274]. At the same time, as discussed in Ghoshray's work, employee surveillance and employee's subjective expectation of privacy shares a competing relationship, and the sheer perception of being "surveilled" can influence one's self-disclosure behavior on the platform [236, 321, 623]. Further, employee's own mental models about LinkedIn privacy might be a factor behind what they share [100].

In summary, `LibRA` is based on self-presentation on a professional social media site, Linkedin. As such, it is subject to variability in self-presentation and motivation found in the population, such as differences in organizational behavior, differences in job status (e.g. looking for a new job vs remaining established), differences in values (e.g. "fixed" vs "growth" mindset), differences in the context of the job (e.g. a software engineer at a small firm vs a large firm) and the assigned job title, and differences in how individuals perceive the positives and drawbacks of their professional information in a publicly accessible space.

These differences should be considered when applying `LibRA` in assessing role ambiguity.

4.3.5 Discussion

Our findings align with the propositions put forth by role theory, that greater `LibRA` measure is associated with factors that are related to depleted wellbeing such as, increased heart rate, increased arousal, decreased sleep, and decreased work hours, and is associated with lower job performance such as decreased task performance and decreased organizational citizenship behavior. Our work bears theoretical, practical, and design implications that surround this new measure of role ambiguity assessed from people's professional social networking data, from the perspective of employees, organizations, and social computing platforms. Our research contributes to the growing interest on the topic of "Future of Work at the Human-Technology Frontier"[4], wherein this study presents new technology-facilitated means to improve workplace "health", performance, and functioning.

Theoretical Implications

This work measures role ambiguity (`LibRA`) for information workers with a diverse intrinsic differences using their self-described portfolios as shared on professional social networking website (LinkedIn). Traditionally, registries and census organizations have served as analogous source of data for people's professional portfolios. This study reveals the feasibility of measuring a role related construct (here `LibRA`) at scale via a previously unexplored, low-cost, and unobtrusive source of data. Management and economics research is advancing in ways that can use this data to operationalize and derive existing measures in novel ways. Thereby, this study revisits old questions in labor economics where existing efforts have been limited to statistical numbers such as salary distribution, unemployment rates, and so on. This study can potentially complement these numbers with richer information on satisfaction and wellbeing at scale.

[4] *https://www.nsf.gov/eng/futureofwork.jsp*

This work lays the foundation of studying employee wellbeing through unobtrusive online data sources that set up marketplaces for employees. These include other professional networking websites such as Meetup, Xing, Jobcase, etc. Being platform-agnostic, methods in this study can be easy to replicate in other platforms or other contexts. In addition, this study combines organizational sciences, and can further be used to advance our understanding of coping mechanisms, incentives, and job satisfaction in general at workplaces, by adopting a technology-focused and technology-driven lens.

Because this work uses individuals' self-described portfolios of job roles and responsibilities, it enables to objectively assess the differences in "what the individual considers and self-describes themselves to be doing", and "what the company hired them for, or what their job description states". That is, the individual may only be showing normative and socially influenced behavior at their work, or show there is information gap, or reveal they intend to invest more effort to learn and gather experience themselves. These behaviors are detectable oblivious to the presence of role ambiguity. Such "unaware role ambiguities" are challenging to capture using traditional approaches as they are tuned to measure the "perceived role ambiguity". Language can reflect differences in personal traits as well as situational ones [250]. This additionally makes our measure less subjectively biased than traditional methods of measuring role ambiguity.

Practical and Technological Implications

Individual-Centric Implications. This study can be used to develop self-reflection tools for employees to mitigate their intrinsic bias in perceived role ambiguity. This can help them continually assess themselves on their skillset and productivity at work. Such self-reflection tools can include within and across organizational role comparisons. These can benefit the employees to have more streamlined information reducing their job search costs and effort, and enhancing their wellbeing. In addition, describing tasks or job role is partially a self-reflection process, and a tool that scores for a type of description will help individuals to identify sources of their role ambiguity.

Further, self-reflection is known to have a positive influence on job satisfaction and wellbeing [174]. Integrating self-reflection tools with our approach would facilitate automated (self)-assessment of one's skillset, interest, and adaptability to an organizational role, and indirectly help them estimate their productivity, wellbeing, and job satisfaction at both their current as well as a future potential workplace. By logging roles, responsibilities, and tasks in a longitudinal fashion, an individual can assess their professional growth and development, and can also be prompted with recommendations for skill training wherever necessary. For the individuals who want to seek professional career-related advice, these logs can function as a diary-style data source to professional mentors and career counselors for better understanding of one's career trajectory, beyond the information presented in a resume.

Organization-Centric Implications. Presently job and skillset training at organizations is not streamlined [442]. Either they train a lot of employees in a batch, or they mentor them individually. However, with more information regarding how employees perceive their role, employers can identify the area of training required that will reduce role ambiguity and enhance the productivity of employees. This method can help reduce the time to identify such role ambiguity gaps, reducing training and employee wellbeing costs. This in turn, can improve employee retention for companies by identifying turnover intentions.

Aligning with and confirming the literature [361], our findings suggest that `LibRA` is not dependent on individual differences such as personality, gender, supervisory role, and executive function. This can inform organizations how these roles or titles can be transformed to match skill-level, task-assignment level, and incentive-level restructuring. This study calls for more careful development of job descriptions. Organizations can involve team- or sector- level staff in curating job descriptions that are more attuned to employees, and can dynamically update the descriptions in accordance with the necessity [312, 349].

The interest in human resource management is still nascent but promising in the HCI and design community. In fact, cross-disciplinary literature pertaining to workplaces and online technologies provide potential use-cases urging the attention of designers [574]. Our work has implications towards designing and developing organization-centric technologies:

(1) First, tools can be built that suggest carefully chosen, fine-tuned job titles to companies, based on `LibRA` [41, 258]. This is particularly important because younger organizations sometimes offer (higher ranking or impressive-sounding) titles to employees in lieu of higher salaries, but this strategy has been reported to backfire due to increased role ambiguity, affecting employee productivity and wellbeing [551]. Adoption of tools that inform organizations about existing ambiguities in specific job roles, therefore, has the potential to make the workplace and individual roles more conducive to effective coping against workplace stress [365]. Moreover, professional social networking platforms (such as LinkedIn) are already heavily used to recruit by job agencies and resume matching consultancies [351]. Such agencies can use insights gained from our approach to match and recommend suitable jobs to prospective employees.

(2) Second, this study can help design workplace tools and dashboards to enhance organization "health" or functioning. Such dashboards can unobtrusively and proactively assess employee role ambiguities at scale, taking employees' privacy considerations into account. In fact, many companies already provide their employees with internal social media platforms [177], online engagement forums, or even email profile description spaces, where they can regularly update their self-explained expertise and role descriptions, along with manager or peer-appraised testimonials. By leveraging such internal datasets, companies can potentially adopt these dashboards to gauge role ambiguity to make informed role matching for open positions in internal hiring. Companies can also restructure and reassign current employees with appropriate incentivization and compensation on their task and workload.

This study showed that *role ambiguity may not necessarily be "bad"*. It is possible that individuals who demonstrate desirable organizational characteristics, such as proactivity and initiative [143], may show high role ambiguity. Therefore, it requires caution in how `LibRA` is made actionable by companies, especially in the light of the many possibilities to build the above organization-centric technologies. Again, companies should not only encourage and provide rewards for these type of employees because they bring role and skill diversity to the organizational culture, but also consider shepherding these individuals with better coping strategies so that they deal better with their wellbeing concerns that are attributed to an underlying role ambiguity [359].

Social Computing Implications

This study bears implications for *social computing system* design. Platforms such as LinkedIn cater to both individuals by recommending them jobs, and to companies by recommending them individuals. Their recommendation algorithms can directly leverage `LibRA`, complementing and going beyond general skills and experience matching. In addition, social computing platforms can aggregate role ambiguities across and within organizations. This will add more transparency on company experiences, complementing review websites (such as Glassdoor) [155]. Finally, LinkedIn already enables individuals to gauge their "professional value" based on their profile stats [633]. An added feature to that could be a measure like `LibRA`, and guided recommendations on the basis of one's weaknesses (in terms of role ambiguity) to online training (such as Lynda[5]), or with classes at local third-party training centers. For privacy-preserving purposes, LinkedIn anonymizes one's list of followers [633], but this also compounds the fact that there is no structured way to know "who sees what on Linkedin", adding complexity in terms of the audience is a problem that an individual faces [305, 389, 403]. However, the platform can adopt design changes such as allowing individuals with diversified interests to create multiple professional personas for different audiences. For example, someone who is both a Software Engineer as well as a part-time Physics Tutor, may self-promote their expertise and gain visibility in both the disciplines but to different controlled audiences [337], who can assess their role-related constructs only

[5] *linkedin.com/company/lynda-com/*, Accessed 2019-03-21

on the discipline that they are interested in.

Ethical, Privacy, Social, and Policy Implications

Back in 2014, when Zhang, De Choudhury, and Grudin studied "creepiness" and privacy concerns related to social media use by workplace professionals, they found concerns shifting from boundary regulation to behavior tracking by social media platforms [682]. One data-driven behavioral inferences has evolved since then, and also come under scrutiny for privacy breaches such as the Cambridge Analytica scandal [94]. This work renews attention to the challenges that may arise when employee data is appropriated for workplace surveillance; as Van Dijck's research noted, "LinkedIn's functionality goes beyond its self-claimed ambition as a professional matchmaker, and ventures into behavioral monitoring." In fact, with research like this that uses people's online self-presentation to infer offline behavior (with high-risk decision outcomes such as career) augments several complexities to the perception of ethics and privacy, and consequently people's social media behavior.

Although this work leverages public social media of individuals, it raises questions on the *privacy-breach of individual information*. An employee's motivations and expectations for LinkedIn might have been only to network or to browse jobs, and they may be well unaware that their published portfolio may also be used to analyze their present or future role-ambiguity and measures of organizational fit or job security [25, 152, 633, 682], which the individuals may not feel comfortable about, especially when this information is made accessible to their employers. *This work is not intended to facilitate employer surveillance*, which shares a competing relationship with employee's subjective expectation of privacy [144, 236, 588].

More elaborately, as per Goffman's theory of self-presentation, individuals may present two kinds of information — one that they intend to "give off", and one that "leaks through" without any intention [249, 421]. This implies that both the perspectives may be present in our sort of research. Publishing role descriptions as online portfolios on social media benefits the individuals in many ways, but methods like this may abuse the data without their consent or awareness. Employers may make inferences about role ambiguity and job satisfaction to make decisions on rewarding, promotions, or even retention and layoffs. To regulate such practices via the use of social media data, employee right protection agencies and lawmaking bodies should consider making guidelines on how organizations engage in data-driven decision making regarding their workforce [321]. This work calls for constitutional jurisprudence in employee social media rights and employer surveillance [236, 321].

Additionally, companies have varying kinds of expectations, history, culture, structure, and needs in their organizations that are latent and beyond what role descriptions say [215, 411, 574]. These factors, alongside platform-related and individuals' intrinsic factors that may impact their role ambiguity assessments, should be accounted for before making decisions merely on any sort of data-driven form of inferred role ambiguity.

On the contrary, *individuals may also start gaming the system*, and describe themselves in language that is more attuned with their role descriptions at work to gain professional advantages [633]. Such deceptive behavior calls for action for stakeholders with diverse interests ranging from academia and industry, as this adds complexities, and may even rigorously change the whole social computing ecosystem on LinkedIn compared to how it is used now. The platform may consider bringing moderation of content or user flair/karma (such as on Reddit) to enable that only credible information is shared on the platform. Presently, LinkedIn already includes features such as testimonials and recommendations that may be leveraged to counter such behavior on LinkedIn. However, such measures are biased as well and can cause Matthew Effect (*the rich get richer, and the poor get poorer*) [414], so accounting for them needs additional consideration.

In addition to the above, our work is only able to measure role ambiguities for those who are on LinkedIn, predominantly in white-collar jobs. According to Pew Survey Reports, 25% of U.S. adults are on LinkedIn, and the demographic is skewed towards the younger, urban, and college-graduate individuals [479]. It is likely that only "privileged" individuals can benefit from these kinds of online- or technology- driven measures to advance and positively impact their job outcomes and wellbeing, e.g., via the self-reflection tools as discussed above. Consequently, those who are not on the platform (which could due to their socio-economic conditions, e.g., the vast majority of blue-collar job workers, or by choice), may feel compelled to use it owing to social and professional pressures of being on it. This adds to the complexities related to *digital inequalities in job-seeking and job summarization behavior on the internet* [259, 328, 685].

CHAPTER 5
ADDRESSING CHALLENGES OF SOCIAL MEDIA DATA BY COMPLEMENTARY MULTIMODAL SENSING DATA

While the previous two chapters show the potential of social media as a passive sensor in human behavior and wellbeing studies, we also need to realize that social media data comes with limitations. This chapter aims at proposing methodologies to overcome some of these limitations by complementing social media with multimodal sensing data.

In the first study, we build machine learning techniques that use social media data as a passive, unobtrusive sensor for inferring mood instability, alongside actively sensed data given by EMAs. We first develop a seed classifier that uses EMAs from a mobile sensing study (CampusLife) as ground truth data to predict binary mood instability status of individuals from their Facebook content. We augment this classifier to improve its robustness and reduce overfitting by adding public data samples from Twitter. We evaluate this augmented classifier on unseen populations of Twitter users who self-disclose suffering from bipolar and borderline personality disorders. This study advances the health sensing research agenda by introducing a new modality of pre-existent, large-scale sensor data—social media, which can significantly improve the modeling and inferential capabilities of small-scale active sensing frameworks.

In the second study, propose a statistical framework to leverage the potential of social media in sensing studies of human behavior, while navigating the challenges associated with its sparsity. Our framework includes principled feature transformation and machine learning models that predict latent social media features from the other passive sensors. We demonstrate the efficacy of this imputation framework via a high correlation of 0.78 between actual and imputed social media features. With the imputed features, we test and validate predictions on psychological constructs like personality traits and affect. We find that adding the social media data streams, in their imputed form, improves the prediction of these measures. This framework can be valuable in multimodal sensing studies that aim to gather comprehensive signals about an individual's state or situation.

5.1 Leveraging EMAs for Groundtruth to Infer Mental Health on Social Media

In the assessment of mental wellbeing, two central constructs are mood and emotion. A number of mood disorders are characterized by patterns of persistence or fluctuation in mood, for example, bipolar disorder, in which the person experiences swings between depressed and elevated mod [22].In fact, for many psychotic disorders and experiences, the ebb and flow of symptoms varies with changes in mood and affect [81, 401, 465].

In addition to being a core criterion for many mood disorders, certain daily mood patterns may also be important for predicting the onset of mood and psychiatric disorders. For instance, a number of studies have observed that individuals who show abnormally unstable moods are more likely to later develop severe psychosis [140]. Specifically, mood instability may indicate challenges with regulating emotions. Emotion regulation is the process, which influences the emotional experience of an individual, as well as, how and when they express an emotion [267]. In theory, the root of most major psychopathologies is the difficulty that the people have when attempting to regulate in terms of theri intensity or duration [488]. Consequently, measuring *mood instability*, or frequent temporal changes of mood, in terms of its two dimensions, *valence* and *arousal* [487], is recognized to be critical in understanding any causal pathways between mood and mental well-being, as well as in developing intervention capabilities that can bring timely help to those in need [487].

Current capabilities to measure mood instability are limited — psychologists and clinicians have deployed survey instruments, such as the Affective Lability Scale [289]. When these instruments ask people to summarize their emotional experiences from a long segment of time in the past, the data can be distorted by recall bias and by bias in the process of interpreting and integrating past experience [577]. When researchers measure affect infrequently, they may further not capture short-term dynamics in mood or the context of the experience, both of which are needed to fully describe the persistence or instability of mood. Taken together, these weaknesses can substantially limit the utility of these instruments for the assessment of mood instability.

Since its emergence in the 1970s, a technique known as "ecological momentary assessment" (EMA), has been increasingly applied to overcome these challenges in questionnaire-based approaches to affect measurement [146, 295, 577]. With EMAs, participants are prompted to respond to survey items sporadically throughout the day as they engage in typical activities. In fact, in recent work, mobile phone applications have been built to make EMA data easier to collect and less burdensome for participants [485, 683]. These modern EMA applications can therefore be considered "active" sensors, in that they require active participation by the individual. While EMA as a form of active sensing enables capturing affective states in an individual's natural habitat and uses a direct method to gather accurate in-the-moment affective information, it requires careful and highly engineered study design, as well as continual, proactive engagement of the user to answer questions [122]. Therefore, it may be vulnerable to high participant burden and may result in low compliance when data acquisition is required for extended periods of time [485, 683]. Researchers have begun to employ various forms of passive sensing [579], such as by logging an individual's phone usage and via wearable sensors, to address these limitations [7, 98, 291, 360, 363, 382]. There has been significant success in these sensing techniques when applied in the context of affect and mood measurement [187, 269, 650]. However, despite the dense, high fidelity data they capture, existing active and passive sensing paradigms are prone to biases and scalability issues due to resource and logistical constraints, such as cost and active compliance of the participants [565, 620].

This article introduces a new modality of passively sensed health, social and behavioral data, specifically that gathered from social media, to overcome some of the challenges noted above. A growing body of work has employed social media data as a "sensor" to identify markers and assess risk to a variety of different health and well-being concerns that have social underpinnings, including mood and affective disorders [136, 166]. In the context of this paper, social media based sensing of moods and their fluctuations over time can capture affective experiences and behaviors spontaneously, reducing the significant bias impacting affect and mood recognition in controlled environments [161]. Moreover, social media data, through quantification of language can enable capturing rich contextual information about mood and its dynamics. However, since sensed data gathered from social media is often sparse and often does not include gold standard markers of well-being states, research has begun to utilize it in conjunction with other conventional forms of sensing, such as active sensing [370]. Our work in this paper extends these early investigations.

Our research objective in this paper examines *whether and how high fidelity active sensing data may be augmented with large-scale, naturalistically-shared social media data to infer mood instability*. The computational investigations presented in this paper leverage a pilot mobile sensing study within the CampusLife project at Georgia Tech, that provided access to 1,606 mobile EMAs over five weeks, and a Facebook archive of 13,340 posts from 23 college student participants. This study also considers a complementary population experiencing a set of mental health challenges who can highly benefit from capabilities that enable sensing mood instability, and who, per literature [22, 141], are likely to exhibit signs of high mood instability. I employ a Twitter corpus of over 21 million posts from 9,654 individuals who self-reported their diagnosis of bipolar or borderline personality disorder on the platform. Using a theoretically-grounded quantification of mood instability from EMAs [323, 622], this study makes the following contributions:

- A seed classifier to detect binary mood instability status (low, high), utilizing the EMA responses of the CampusLife participants as ground truth, and psycholinguistic attributes from their Facebook posts as features.

- A semi-supervised machine learning framework to augment the above seed classifier of mood instability by incorporating data samples acquired from Twitter; specifically,

Table 5.1: Descriptive statistics of the DASS-21 data collected through enrollment questionnaire. Levels inferred per prior work [253].

Level	Dep.	Anx.	Str.
Normal	26	27	26
Mild	5	6	9
Moderate	8	10	8
Severe	3	0	1
Ex. Severe	3	2	1

Table 5.2: Descriptive statistics of the Facebook seed dataset collected from 23 participants in the CampusLife study.

Feature	count	mean	median	stdev.
Friends	10,578	459.91	372	321.20
Likes	3358	152.64	102	173.22
Profile Pictures	433	18.83	9	18.00
Status	13,340	580.00	294	713.48

an approach by which the model can learn from both (scarce) labeled and (voluminous) unlabeled data around mood instability.

- A lexicon of language cues appropriated on Twitter, that are highly indicative of low or high mood instability.

5.1.1 Study and Data

The sensing data employed in this work is derived from a larger mobile sensing study that was conducted in April 2016 involving several college students at Georgia Tech, a large public university in the southeast of the U.S. The study was approved by the Institutional Review Board at Georgia Tech (#H16009).

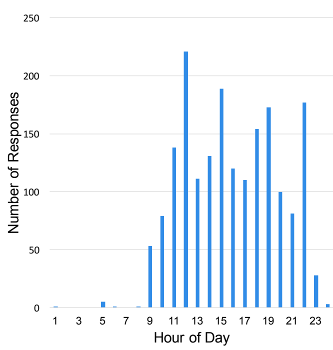
Participants Participants (undergraduate and graduate students at the university) were recruited by word of mouth, flyers, and social media advertising. In addition, recruitment email messages were sent to students by the registrar and by instructors of a mandatory course for undergraduates. A total of 51 participants enrolled in a five week-long study (40% females and 60% males; 46% undergraduates and 54% graduates; mean age 22 years). Similar to the study conducted within the StudentLife project [651], the smartphones (the Android operating system only) of the participants were instrumented to collect a variety of actively and passively sensed data: active data were collected through a commercial EMA platform, called Quedget and passive data were collected through the sensors on board. Participants also answered different psychological survey questionnaires, during the study.

Table 5.3: Descriptive statistics of the EMA data collected in the CampusLife study.

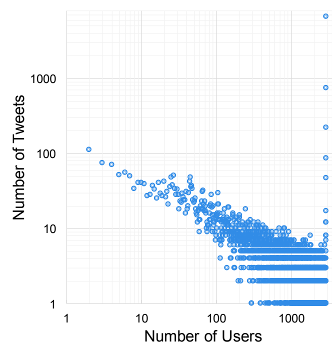
Metric	Value
Number of Participants	51
Number of Responses	1,606
Mean of Responses/Participant	31.49
Median of Responses/Participant	28.00
StDev. of Responses/Participant	21.13
Period of study	5 wks.

Table 5.4: Descriptive statistics of the Twitter mental health dataset.

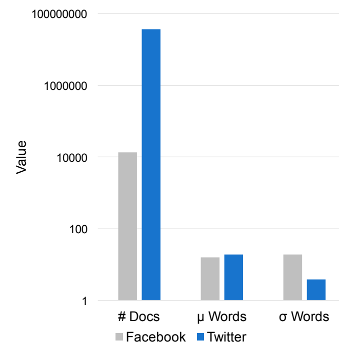
Metric	Borderline	Bipolar	Control
Number of Users	6,326	3,328	9,394
Number of Tweets	14,780,813	7,095,801	15,136,451
Number of Tokens	194,801,582	101,397,309	411,656,658
Mean of Tweets/User	2,336.52	2,132.15	1,611.29
Median of Tweets/User	3,398.00	2,858.00	1,131.00
StDev. of Tweets/User	1,310.27	1,374.35	1,432.05



(a) Distribution of EMA responses by the hour of the day.



(b) Distribution of number of users by number of tweets in the mental health dataset.



(c) Comparative statistics of Facebook seed dataset and Twitter mental health dataset.

At the end of the study, participants could also volunteer for a one-time access to their social media data.

Study Procedure The CampusLife study consisted of orientation, data collection and exit stages, as also employed in the StudentLife study [651].

Orientation. At the start of the study (orientation), participants were first required to watch a pre-recorded video and tutorial developed by the study team that described the research goals of the study, the study procedures, the types and mechanisms of data collection, the privacy considerations, as well as the risks and benefits involved in participation. Each participant was then provided with an IRB approved consent form to sign; on signing this form, participants agreed to allow the research team to acquire active and passive sensing data from their personal Android smartphones. Members of the research team also familiarized the participants with the EMA software interface and the procedure for responding to EMA prompts on their phones.

During the study, participants' intent to share social media data was sought through a yes/no survey question. The participants were then directed to an online survey (administered through Qualtrics) that included a battery of already validated questionnaires to assess their mood, individual differences (e.g., demographics), and mental well-being status (Perceived Stress Scale (PSS) , Flourishing Scale, and Depression Anxiety and Stress Scale (DASS-21) [24]). The purpose of using these questionnaires was to establish baselines for the mental well-being of the participants: For instance, among the participants who answered the DASS-21 questionnaire, we found notable variation in their psychological well-being, although our study population is not a clinical one. That is, mapping the responses given on the DASS-21 scale to levels of depression, anxiety, and stress based on prior work [253], we observe that about 47% of our participants showed above-normal levels of either depression, anxiety or stress (Table 5.1). This indicates the presence of sufficient proportion of the population in which a wide range of expression of mood instability levels may be expected.

Exit Stage. At the time of study conclusion (exit stage), participants who indicated affirmatively about their willingness to share social media data during orientation, met with the researchers face to face. During this meeting, they were provided with a second consent form for the social media data collection. This consent form enabled a one-time download of a participant's social media, specifically Facebook and/or Twitter archives which we then de-identified and stored in a secure, encrypted server for the ensuing analysis presented in this paper. Participants could consent to share both or either of these two types of social media data. All participants were instructed to bring their laptop computers to this meeting, as a privacy-preserving mechanism to download their Facebook data. Our choice of Facebook and Twitter as the two social media platforms was driven by statistics from the Pew Research Center [262], which reports these as two of the most popular platforms in college-aged populations.

Incentives. Participants were given $40 for the time and effort required to enroll in the study and install software on their phones. Additionally, they were paid in direct proportion for each answered EMA (up to a maximum of $40). They were given additional compensation if they consented to share any of their social media data, with a maximum of $40 if they shared all of the social media data we requested. As an incentive to attend the information sessions in which we introduced the study, we provided food, beverages, and $5 gift cards to those in attendance.

Sensing Data

EMA Data As mentioned above, EMA data was collected from the participants through a platform known as Quedget. The process of responding to a questionnaire item can impose a burden even before a subject faces the labor of responding. Quedget is designed to use the lock-screen of a smartphone as a way of gaining attention only when a subject is between operations. On this Android-only platform, a researcher-defined schedule determines when and which questionnaire item is displayed on the lock-screen of a participant's phone. This study defined four mutually exclusive four-hour long time windows between 9 am and 11 pm, during participants were presented with questions. Within a specific window, Quedget calculated a random time to trigger the prompt.

The EMAs spanned a variety of questions, such as a Photographic Affect Meter (PAM) [485] (see Figure 5.2). Psychology literature situates valence and arousal dimensions to comprehensively describe an individual's affective state at any moment [488], and PAM has been found to be well-suited for the purpose [485]. The PAM EMA questions showed participants a set of images ordered in a 4 $\times$4 grid, where each image corresponded to a mood of specific valence and arousal (e.g., "angry", "excited", "satisfied", "tired") – the rightmost top image in the grid refer to High Valence and High Arousal, whereas the leftmost bottom image refer to Low Valence and Low Arousal (see Figure 5.2). Participants could select the image that best captured their current mood. By the end of the five-week study period, a total of 1,606 PAM valid EMA responses were collected, spanning all of the

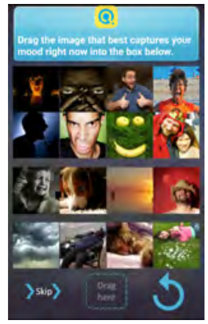


Figure 5.2: Screenshot of PAM.

participants. The participants responded to these EMA questions mostly during the period of 9AM to 10PM, based on the previously described Quedget schedule (ref: Figure 5.1a). Table 5.3 reports the descriptive statistics of the PAM EMA data.

Social Media Data At the exit stage of the study, participants who consented in sharing their Facebook data, downloaded and shared their Facebook account's data dump (starting from the data of their account creation to the date of data collection) as HTML files on laptops. To create data dumps, the participants used a feature provided by Facebook. Participants were instructed to personally delete their private messages and all photos from the created data dumps. The downloaded and curated data from the participants' laptops was then stored by one of the study support volunteers into a detachable hard drive (per the approved IRB instructions), which was eventually uploaded to a secure, encrypted server. These downloaded Facebook data files were parsed, and timeline activities were extracted (which includes "like" information, friend connections initiated and accepted, posts about profile pictures, status updates, check ins into different locations, etc.). This study mainly focuses on the linguistic content of the 13,340 posts of the participants. Table 5.2 reports the statistics of the Facebook data. Out of the 51 participants who provided EMA data, 23 participants provided EMA data, provided their Facebook data. This is referred to as the dataset of 23 participants in the rest of this paper as the "seed dataset".

Also, 10 participants in the above population of 23 also consented to share public Twitter

Table 5.5: Search phrase/method for Twitter data collection for three different samples.

Bipolar	**Borderline**	**Control**
"i have bipolar disorder"	"i have borderline personality disorder"	Twitter stream with *en* as filter
"i have been diagnosed with bipolar disorder"	"i have been diagnosed with borderline personality disorder"	Remove usernames in Bipolar and Borderline datasets
"i was diagnosed with bipolar disorder"	"i have been diagnosed with bpd"	Remove user timelines with text containing 'bipolar disorder'
	"i suffer from bpd"	Remove user timelines with text containing 'borderline personality disorder'
	"i was diagnosed with bpd"	
	"i am suffering from bpd"	

data; that is, they shared their Twitter usernames (or handles) during the exit stage of the study. Utilizing these handles as query terms, I leveraged Twitter's official API to obtain their posts shared in [1]. This is referred to as the "validation dataset". It contains 1425 posts in all, with a mean and median of 142.5 and 58.5 posts per participant respectively.

Twitter Mental Health Data

One of the limitations of the dataset collected through the above CampusLife mobile sensing study is its small sample size, which may present challenges in building computational models of predicting mood instability. To circumvent this challenge, and to develop inference frameworks for mood instability on larger scale of social media data, we choose Twitter as a source to augment our existing social media data gathered through the CampusLife study. Our choice of Twitter data is motivated by prior work: due to its largely public nature, in contrast to Facebook, Twitter data has been used to study mental health concerns [162, 166]. This facilitates not only the collection of large-scale data toward studies like ours, but also, enables identifying and gathering data of individuals who publicly share self-reported diagnosis of their mental health conditions, such as depression, bipolar disorder, or post-traumatic stress disorder [136]. This study is particularly interested in augmenting our seed data from Facebook with complementary Twitter data of individuals who are likely to exhibit a wide range of mood instability. As noted earlier, two conditions wherein sufferers

[1] The API returns the last 3,200 posts of a given user, which in most cases of an average Twitter user, covers their entire timeline data.

are known to be challenged by significant mood variability include bipolar disorder and borderline personality disorder [22, 141, 465]. Our Twitter mental health data collection pursued a strategy to collect data around these two conditions.

This second data collection is conducted by searching for tweets with the Twitter Search API, wherein users had made explicit self-disclosure of the diagnosis or experience of bipolar disorder and borderline personality disorder. These searches were spawned with a set of keyphrases given in Table 5.5. The choice of the keyphrases were motivated from prior work where a similar data acquisition strategy has been successfully applied to identify populations struggling with a mental illness [136], and where it has been observed that these self-reports do indeed capture actual clinical conditions as assessed by experts and psychiatrists [169]. Next, for all of these tweets, I query the timelines of their authors using the Twitter API. Each user timeline refers to a collection of text (capped to a maximum of 3,200) tweeted by a single user. Using this mechanism, we collected 6,326 and 3,328 user timelines of individuals who disclosed the diagnosis or experience of bipolar disorder and borderline personality disorder respectively. Hereforth, we refer to these datasets as Bipolar and Borderline respectively. I also collect an independent sample of tweets using the Twitter Streaming API, which returns live tweets at a particular time. I repeat the above approach to fetch user timelines for these tweets; then I filter out any user who occurr in the datasets Bipolar or Borderline, or if they mentioned 'bipolar disorder' or 'borderline disorder' in their tweets. This third sample of filtered user timelines results to 9,394 users and I call this dataset Control in this paper.

The Twitter mental health dataset finally comprises a total of over 37 million tweets shared by 19,048 unique users (ref: Fig Figure 5.1b gives the tweet to user distribution). Table 5.4 reports the descriptive statistics of this dataset, and Figure 5.1c shows a comparative plot of the seed (Facebook) and mental health (Twitter) datasets.

5.1.2 Method

Quantifying Mood Instability

The participants logged their mood via a set of 16 distinct PAM images, arranged in a 4×4 grid, where valence and arousal increase along the horizontal and vertical axes respectively. I refer to the literature on PAM [485] to map the 16 PAM moods into numeric tuples of valence and arousal values— these values are derived from the absolute position of a mood image in the 4×4 grid, where the values can be -2, -1, 1 and 2 [485]. Using the mapping given in Table 5.6, I quantify a participant's momentary mood states in terms of valence and arousal, which I next use to quantify their mood instability.

To quantify mood instability of a participant, it is necessary to calculate the successive differences in momentary mood states logged by the participants. Since the consecutive observations for any participant do not have uniform time differences in our study (EMAs were randomly triggered at different times of the day), changes or fluctuations in mood cannot be quantified from simple time series analysis of EMA responses. Therefore, I adopt a method proposed in [323] to compute the *Adjusted Successive Difference* (ASD) functions for the valence (and arousal) dimensions of a participant's mood. If x_i is the valence (or arousal) of a participant's logged mood state at time t_i, I compute its ASDs based on Equations Equation 5.1 and Equation 5.2:

$$ASD_{i+1} = \frac{x_{i+1} - x_i}{[(t_{i+1} - t_i)/Mdn(t_{i+1} - t_i)]^{\lambda}} \tag{5.1}$$

Here λ is chosen by minimizing the following cost function, sum of square of the error

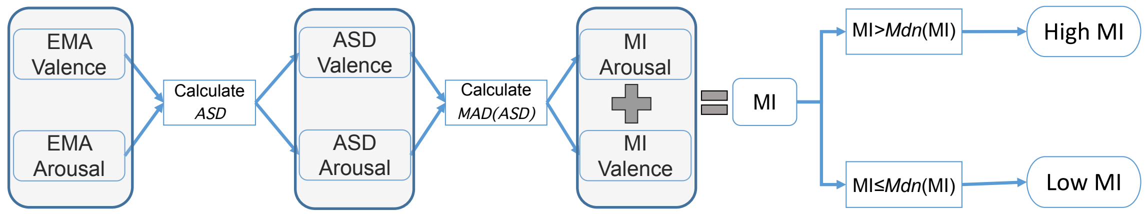


Figure 5.3: A schematic diagram showing the computation of the HighMI and LowMI classes with EMA data.

of expectation $(SSEE)$:

$$\begin{aligned} SSEE(\lambda) &= \sum_i [EAASD_{(t_{i+1}-t_i)}(\lambda) - C(\lambda)]^2 \\ &= \sum_{i=1}^{N-1} \left\{ E\left\{ \frac{|x_{i+1}-x_i|}{[(t_{i+1}-t_i/Mdn(t_{i+1}-t_i)]} \right\} - C(\lambda) \right\}^2 \end{aligned} \tag{5.2}$$

The expected absolute successive difference $(EASD)$ is obtained by nonparametric curve fitting regression method of lowess—a method for fitting a smooth curve [127]. Further, the expected adjusted absolute successive difference $(EAASD)$ is calculated by an adjustment, which eliminates the dependency of $EASD$ on the time intervals. The $EAASD(\lambda)$ at the median time interval is used as the C(λ) (Equation 5.2).

Once I obtain the valence (and arousal) ASD of all participant's mood states reported throughout the study period, I calculate the mean absolute deviation (or MAD) for each of the ASDs, referred to as $MAD(ASD_v)$ and $MAD(ASD_a)$, corresponding to the valence and arousal dimensions respectively.

The sum of $MAD(ASD_v)$ and $MAD(ASD_a)$ is then referred to as a participant's overall mood instability MI throughout the study period — a high MI would indicate that either valence or arousal or both dimensions of their mood states tend to generally show large fluctuations, whereas lower values of MI would imply one or both of the dimensions to exhibit fewer shifts over time. Finally, employing the median of the MI distribution over all participants as a threshold, I categorize the participants into binary class labels, HighMI and LowMI respectively. Those with MI above the median of the MI distribution are assigned the HighMI class, and those with MI under the median are assigned the LowMI class. The steps involved in categorizing users as HighMI and LowMI from EMA data, in Figure 5.3, and with the following equations:

$$MI = MAD(ASD_v) + MAD(ASD_a)$$
$$MI \text{ Class Label} = \begin{cases} \mathsf{LowMI} & \text{if } MI \leq Mdn(MI) \\ \mathsf{HighMI} & \text{otherwise} \end{cases} \tag{5.3}$$

Note that median is a conservative, yet intrinsically understandable and robust measure for central tendency of a distribution. Hence it is adopted as a decision boundary for assessing levels of mood instability in the participants of our study. Although a continuous estimate of the distribution would have been a better quantification of mood instability, it would have made the MI inference task far more difficult, especially in this case, with only a small amount of ground truth data.

Building a Seed Classifier of Mood Instability

Utilizing the above inferred binary levels of mood instability (HighMI and LowMI) in the participants of the CampusLife study, I build a classification framework to predict these class labels from the seed dataset. Although the CampusLife study also acquired Twitter data from a small set of the participants, Facebook is employed as the data source for the seed classifier as it provides with a larger sample of ground-truth labels over Twitter (23 vs. 10 participants).

To build a classifier for mood instability, I extract psycholinguistic features from participants' Facebook data. I focus on the status messages shared on their timeline. I employ Linguistic Inquiry and Word Count, or LIWC [475] on the Facebook posts—this psycholinguistic lexicon has been extensively applied and validated on several studies of social media, behaviors, moods, and mental health [162, 166]. I use 50 of the most relevant LIWC categories per prior work [166], grouped as: (1) *affective attributes* (categories: anger, anxiety, sadness, swear, positive and negative affect), (2) *cognitive attributes* (categories: cognitive mech, discrepancies, inhibition, negation, causation, certainty, and tentativeness), (3) *temporal references* (categories: future, past and present tense), (4) *interpersonal focus* (categories: first person singular pronoun, second person plural pronoun, third person plural pronoun and indefinite pronoun) (5) *lexical density and awareness* (categories: adverbs, verbs, exclusive, inclusive and preposition), (6) *perception* (categories: feel, insight, percept and see), and (7) *social/personal concerns* (categories: achievement, bio, body, death, home, humans, sexual and social). For every participant, I aggregate the occurrence of the word and word stems in each of these LIWC categories, followed by their normalization based on the total number of tokens (words) in the participants' posts. Using this approach, I construct a feature vector of 50 dimensions, for the participants.

I build a supervised machine learning models using the data obtained so far in this section—the ground truth labels of mood instability (HighMI and LowMI) in the 23 CampusLife study participants (dependent variable), and the psycholinguistic features extracted with the LIWC lexicon above (independent variables). I evaluate multiple classifiers, including Naive Bayes, Logistic Regression, Random Forest and Support Vector Machines (with different kernels such as linear, radial basis functions and polynomial). I employ a k-fold cross validation (k=5) strategy for parameter tuning.

Semi-Supervised Modeling of Mood Instability

Note that the number of examples in the seed training data from Facebook (23) is much smaller than the dimensionality of the feature set (50), which risks the seed classifier C_0 in overfitting the data. Semi-supervised learning is a recommended technique in cases where labeled data is expensive or scarce, but where unlabeled data is abundant and significantly

Table 5.6: Mapping of PAM categories to numeric values of Valence and Arousal as per prior work [485].

PAM	Valence	Arousal
Afraid	-2	2
Angry	-1	1
Calm	1	-1
Delighted	2	2
Excited	1	2
Frustrated	-2	1
Glad	2	1
Gloomy	-2	-2
Happy	1	1
Miserable	-2	-1
Sad	-1	-1
Satisfied	2	-1
Serene	2	-2
Sleepy	1	-2
Tense	-1	2
Tired	-1	-2

Algorithm 1: Semi Supervised Mood Instability Classifier

Input: CampusLife Facebook Data F (Seed Dataset), Twitter User Timelines T (Target Datasets).
Output: Mood Instability MI of Twitter Users
X_0, $Y_0 \leftarrow$ Psycholinguistic Features, Mood Instability of F
$T_1, T_2 \leftarrow$ Random Samples of T $\{T_1 < T_2\}$
$X_1 \leftarrow$ Psycholinguistic Features of T_1
$X_2 \leftarrow$ Psycholinguistic Features of T_2
Classifier $C_0 \leftarrow$ SVM (X_0, Y_0)
Clusters $< S > \leftarrow K$-Means Clustering (X_1)
$CD \leftarrow$ Initialize Dictionary $< Key, Value >$
for *every i in K* **do**
    $cc[i] \leftarrow$ computeVectorCentroid ($S[i]$)
    $l[i] \leftarrow C_0$.predict ($cc[i]$)
    Add $< cc[i], l[i] >$ as $< Key, Value >$ in CD
end
for *every i in length(X_1)* **do**
    $label \leftarrow$ Value for $S[i]$ in CD
    Add $label$ to Y_1
end
$X \leftarrow$ concatenate (X_0 + X_1)
$Y \leftarrow$ concatenate (Y_0 + Y_1)
Classifier $C \leftarrow$ SVM (X, Y)
$Y_2 \leftarrow C$.predict (X_2)
return $Y_1 + Y_2$

easy to gather [115]. Unlike completely supervised learning such as classification, these approaches devise ways of utilizing *both* labeled and unlabeled data to learn better models. In prior work, similar methods have also been used in problem domains where positive examples are a considerably rare occurrence, creating huge imbalance between the sizes of labeled and unlabeled data [684]. These conditions satisfy our context as well. Thus we employ a semi-supervised approach of improving the robustness of C_0, by augmenting it with training data from the Twitter mental health datasets (Bipolar, Borderline and Control).

Establishing Linguistic Equivalence The above semi-supervised learning approach involves combining datasets spanning multiple social media platforms (Facebook, Twitter) and multiple populations (college students vs. general online population). Therefore, I conduct two tests of linguistic equivalence to demonstrate the feasibility of adopting the semi-supervised learning approach. The tests aim to establish that: a) content shared across the seed and mental health datasets (from Facebook and Twitter respectively) are comparable—establishing *cross-platform equivalence*; and b) that social media data of a college population (the CampusLife participants) may be utilized to measure mood instability in an independent population self-reporting bipolar or BPD diagnoses (Bipolar and Borderline data), and whose specific demographics are unknown—establishing *cross-population equivalence*. I adopt an approach involving pairwise comparison of word vectors, drawing from a similar technique in the computational linguistics literature [34]. The technique involves first constructing word vectors using the frequently occurring n-grams in each source of data, and then employing a distance metric, e.g., cosine similarity, to assess their linguistic similarity. Cosine similarity of word vectors is an effective measure of quantifying the linguistic similarity between two datasets [476], and a high value would indicate that the posts in the two datasets are linguistically equivalent.

To establish cross-platform equivalence, I extract the most frequent 500 n-grams from the seed dataset (Facebook), and the same from the mental health datasets (Twitter) (sample size = 10,000). Next, using the word-vectors of these top n-grams (obtained from the Google News dataset of about 100 billion words [417]), I compute the cosine similarity of the two datasets in a 300-dimensional vector space. We observe that seed and mental health datasets exhibit high cosine similarity (0.9), providing confidence in the use of the semi-supervised learning approach. Additionally, I conduct a pairwise equivalence test to validate the linguistic similarity between the Facebook and Twitter data of the same participants, using the same technique. We do not observe any significant differences in the manner in which Facebook and Twitter are used in our participant pool — for the 10 participants for whom we have both Facebook and Twitter data, we noted high similarity (mean=0.85, std.=0.15) in linguistic attributes (n-grams).

Towards assessing cross-population equivalence, I again employ word vector comparison to first assess if the cosine similarity between the word vectors of the Twitter data of the 10

CampusLife participants and that in the Bipolar and Borderline datasets is high. This leads the similarity to be 0.94 and 0.95 respectively, indicating that the college student participants' social media data is linguistically similar to the unlabeled mental health datasets we use in our ensuing semi-supervised learning approach.

To assess the correspondence between psycholinguistic features from Facebook and Twitter posts of same participants, I conduct two-sample Kolmogorov-Smirnov tests (KS tests) for all LIWC features. The KS-statistic is very low, ranging between 0.01 and 0.38 across the features (median = 0.07 and standard deviation = 0.08), and only 33 out of 50 features exhibit a significance (p<0.05). This suggests that there is very little significant statistical difference between the features of Facebook and Twitter datasets of the 10 participants who shared their data from both the sources.

Augmenting Training Data with Self-Training

Once cross-platform and cross-population linguistic equivalence is established, I proceed with the semi-supervised learning approach. This study borrows from a method known as "self-training" that assumes the data to naturally cluster into groups (in our case we would expect HighMI and LowMI to exhibit similarities in their respective behaviors), and therefore employs a clustering algorithm to categorize the whole dataset, and then label each cluster with labeled data [150].

First, I proportionately separate random samples of 200, 100 and 300 users from our Twitter target datasets, Bipolar, Borderline and Control. Next, I cluster these users in an unsupervised fashion using K-Means (K=2) clustering. For each of these clusters, I find the cluster centroids, and machine label the cluster centroids using C_0. Using the predicted labels of cluster centroids as labels, I augment the training data with 600 additional users from Twitter. I describe our algorithm of classification in algorithm 1.

Machine Labeling of Mood Instability in Unseen Data

I use the augmented dataset of 623 users (23 from the CampusLife study and 600 from Twitter) to build a mood instability classifier C. I extract psycholinguistic features for the posts of each user in this augmented dataset. Since the volume of posts of the users within the Facebook and Twitter (seed and target) datasets vary significantly, I standardize the feature vectors separately for our Facebook and Twitter dataset (i.e., re-scaling each feature distribution to zero mean and unit variance [334]). Like before, I evaluate multiple classification models, and use k-fold cross-validation (k=5). I employ the trained classifier to predict the mood instability labels (HighMI and LowMI) of the users in the held out target datasets, Bipolar, Borderline, and Control respectively.

Characterizing the Language of Inferred Mood Instability

This final subsection presents the methods I use for characterizing the language expressed in social media that relate to HighMI and LowMI. Specifically, on the corpus of the posts of all of the Bipolar, Borderline, and Control users that are labeled or inferred to be of HighMI or LowMI, I extract the top occurring most relevant n-grams (n=1, 2, 3) and compute their Log Likelihood Ratio (LLR) [175] across the two classes HighMI and LowMI. We consider the minimum threshold of occurrence for an n-gram in any class as 500, and then calculate the probability of occurrence of every such n-gram in the HighMI, to the same in the LowMI. The LLR for an n-gram is determined by calculating the logarithm (base 2) of the ratio of its two probabilities, following add-1 smoothing [332]. Thus, when an n-gram is comparably frequent in the two classes, its LLR is close to 0; it is closer to 1, when the n-gram is more frequent in HighMI, whereas, closer to -1, for the converse.

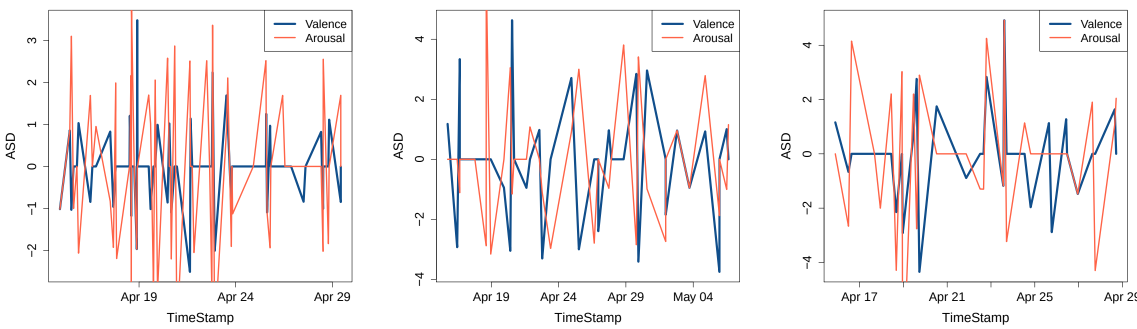


Figure 5.4: Adjusted Successive Difference (ASD) plots of EMAs for three sample participants in the seed dataset.

Table 5.7: Accuracy of the seed mood instability classifier (C_0) based on k-fold cross-validation (k=5) on the seed dataset of 23 CampusLife participants.

Metric	**Mean**	**Stdev.**	**Mdn.**	**Max.**
Naive Bayes	0.58	0.54	0.75	0.83
Logistic Regression	0.51	0.35	0.50	0.80
Random Forest	0.48	0.64	0.50	0.83
SVM (Kernel=Poly.)	0.56	0.24	0.50	0.80
SVM (Kernel=RBF)	0.51	0.35	0.50	0.80
SVM (Kernel=Linear)	0.68	0.29	0.75	0.83

Table 5.8: Augmented training data following K-Means (K=2) clustering.

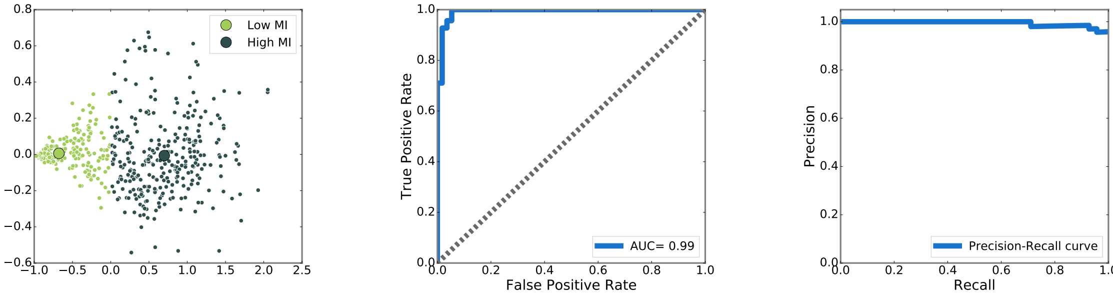
Data	HighMI	LowMI	**Total**
CampusLife	11	12	23
Bipolar	120	80	200
Borderline	65	35	100
Control	110	190	300
Total	306	317	623

5.1.3 Results

Seed Classifier for Mood Instability

Now, I present the results of developing a seed classifier of mood instability, utilizing the Facebook data of the 23 CampusLife participants, and their mood instability labels (HighMI and LowMI) inferred from their EMA data during the study period.

To quantify these mood instability labels, I begin by calculating *Adjusted Successive Difference* (ASD) values of the EMA responses for each of the CampusLife participants. First, I find λ by minimizing cost function, or sum of square of successive differences $(SSEE(\lambda))$ as defined in Equation 5.2. For this purpose, I iterate on $n = [1, 10]$, where $\lambda = 1/n$, chosen based on the method described in [622]. Figure 5.4 shows the ASD curves for three sample participants in the study. Per these ASD values, I find that the overall MI of the participants in the study ranges from 1.65 to 30.8, with a median value of 3.14. Based on the definition



(a) (b) (c)

Figure 5.5: (a) A two-dimensional representation of the k-means clusters. The axes correspond to the two largest principal components. (b) ROC (Receiver Operating Characteristic) curve of mood instability classifier ($\mathtt{C}$), built with augmented training data. (c) Precision-Recall curve of classifier $\mathtt{C}$, built with augmented training data. Small slope indicates good performance.

Table 5.9: Performance metrics of mood instability classification ($\mathtt{C}$) based on k-fold cross-validation (k=5) on seed dataset of 23 CampusLife participants.

Metric	Mean	Stdev.	Mdn.	Max.
Accuracy	0.68	0.29	0.75	0.83
Precision	0.66	0.49	0.83	0.88
Recall	0.68	0.31	0.83	0.83
F1-score	0.64	0.38	0.73	0.83

Table 5.10: Performance metrics of mood instability classification ($\mathtt{C}$) based on k-fold cross-validation (k=5) on the augmented data of 623 users.

Metric	Mean	Stdev.	Mdn.	Max.
Accuracy	0.96	0.09	0.98	0.99
Precision	0.96	0.07	0.98	0.99
Recall	0.96	0.09	0.98	0.99
F1-score	0.96	0.09	0.98	0.99

of HighMI and LowMI given in Figure 5.3, I obtain 11 and 12 users belonging to these two classes respectively. To build a seed classifier for mood instability, I extract frequency of occurrences of the psycholinguistic categories from the above labeled seed dataset of 23 CampusLife participants. After normalizing the distribution of these occurrences, I use them as features and build several classification algorithms on the binary mood instability labels HighMI and LowMI. Table 5.7 summarizes the accuracy returned by each of these classification algorithms, including Naive Bayes, Logistic Regression, Random Forest, and Support Vector Machines (SVM) with different kernels based on k-fold cross-validation ($k = 5$). The SVM Classifier with linear kernel returns the highest accuracy (mean=0.68 and max.=0.83). This motivates the choice for using this as the seed classifier of mood instability. I refer to it as the $\mathtt{C}_0$ model.

Table 5.11: Confusion Matrix of Mood instability classification (C) based on users in unseen data from three different twitter samples.

Data	HighMI	LowMI	**Total**	% HighMI
Bipolar	3863	2232	6095	63.38
Borderline	1997	1208	3205	62.31
Control	3272	5510	8782	37.26

Table 5.12: Results of independent sample t-test comparing target (Bipolar and Borderline) and control datasets for mood instability classification.

Data	t-**stat**	p
Bipolar	32.97	***
Borderline	24.13	***

Classification with Semi-Supervised Learning

Now I present the results of augmenting the above seed classifier of mood instability (C_0) with additional training data from the target datasets (Bipolar, Borderline, and Control); for the purpose, I employ the semi-supervised learning method described in the methods subsection.

To build a semi-supervised mood instability classifier, I apply K-Means (K=2) clustering on LIWC feature vectors from a dataset of 600 Twitter users sampled from Bipolar, Borderline, and Control. The choice of two clusters is motivated from the observation that I intend to identify groups of users exhibiting one of the two mood instability labels—HighMI or LowMI. I obtain two clusters with 295 and 305 user vectors respectively. Figure 5.5a shows a visual 2-dimensional representation of these clusters based on the two largest eigenvectors – I use principal component analysis [330] to extract the eigenvectors of the user vectors in each cluster. I label the cluster centroids using C_0 classifier, to determine the first cluster consists of users with HighMI, and the second consists those with LowMI. These cluster-labeled data, along with the labeled Facebook data of the 23 CampusLife participants (623 users in all) becomes the augmented training data (ref: Table 5.8. This augmented dataset has 306 and 317 users with HighMI and LowMI labels respectively.

Next, with this data, I build multiple classifiers of mood instability, with an SVM classifier C with linear kernel, yielding the best performance described as follows. I obtain an Area under curve (AUC) of 0.99 for C's Receiver operating characteristic (ROC), Figure 5.5b shows the ROC curve of C and Figure 5.5c gives the precision-recall curve. I validate C, on the augmented data obtained above, using a k-fold cross-validation (k=5). The C model gives a mean accuracy of 0.68 and 0.96 on the seed and augmented training data respectively. I report these performance metrics in Table 5.9 and Table 5.10 respectively, for the seed and augmented training datasets. Based on these numbers, I infer that the classifier C is stable and works well on the augmented data containing target datasets from Bipolar, Borderline, and Control, without dropping accuracy in classifying the seed data of the 23 CampusLife participants.

I apply the classifier C on the remaining held out target datasets (6,095 Bipolar users, 3,205 Borderline users, and 8,782 Control users), to machine label them. I report the distribution of the mood instability classifier C across the three target dataset samples in Table 5.11. We observe that HighMI users occur in about 64% (out of 6,095), 62% (out of 3,205) and 37% (out of 8,782) of the users in Bipolar, Borderline and Control data samples respectively. An independent sample t-test of the labeled users each from Bipolar and Borderline, with Control shows statistical significance at the α=0.05/n (n=2) level, following Bonferroni correction (ref: Table 5.12). In other words, these numbers indicate that the likelihood of Twitter users self-reporting diagnoses about bipolar or borderline personality disorders are almost twice as likely to exhibit high mood instability compared to those who do not self-disclose of these conditions.

Validation of the Mood Instability Classifier

In order to validate the performance of the mood instability classifier C, I evaluate its accuracy on an unseen MI labeled dataset of CampusLife participants. For the 10 participants, who shared their public Twitter feeds within the CampusLife study, I infer the mood instability (HighMI and LowMI) using classifier C. Comparing these inferred MI labels with the actual labels of the participants, we observe that C correctly predicts the MI label of 9 of these 10 participants. This affirms the claim that C works satisfactorily across platforms and is able to correctly infer MI in the population of college students based on their social media data.

Table 5.13: Comparison of MI classification in the mental health Twitter datasets using the seed classifier $\mathtt{C}_0$ and the semi-supervised learning based classifier $\mathtt{C}$. The higher standard deviation ($stdev.$) in the distribution of k-fold cross validation (CV) accuracies of classifier $\mathtt{C}_0$ shows its high sensitivity (and therefore instability) across different folds.

Data ↓	**k-fold CV accuracies of $\mathtt{C}_0$ (% HighMI)**							**k-fold CV accuracies of $\mathtt{C}$ (% HighMI)**						
Folds →	**1**	**2**	**3**	**4**	**5**	**Mean**	**Stdev.**	**1**	**2**	**3**	**4**	**5**	**Mean**	**Stdev.**
Bipolar	66.81	69.86	64.64	43.76	62.82	51.38	10.30	62.87	63.64	62.66	63.18	63.38	63.15	0.39
Borderline	61.37	63.81	54.41	34.04	56.13	45.06	11.76	61.06	61.81	62.44	62.84	62.31	62.09	0.68
Control	42.04	46.05	37.35	24.79	37.94	31.40	7.99	36.70	36.54	36.56	36.47	37.26	36.71	0.32

I evaluate how the mood instability classifier $\mathtt{C}$ improves over the performance of the seed classifier $\mathtt{C}_0$. In particular, I compare the decision functions of $\mathtt{C}_0$ and $\mathtt{C}$. A decision function estimates the confidence score of a training sample, based on the distance of the data points from the hyperplane in an SVM classifier [89]. These points are referred to as the support vectors (in a vector space, a point can be thought of as a vector between the origin and that point). In this case, the mean value of the decision function of $\mathtt{C}$ is 94% higher (1.54 vs. 0.79) than that of $\mathtt{C}_0$, showing remarkably higher confidence in model fitting. This suggests that classifier $\mathtt{C}$ performs better than $\mathtt{C}_0$ on an MI labeled dataset in terms of model fit and confidence.

In addition, I compute MI in the unlabeled mental health Twitter datasets, using the $\mathtt{k}$-folds ($\mathtt{k}$=5) of Classifier $\mathtt{C}_0$ and $\mathtt{C}$. We observe that $\mathtt{C}_0$ shows an unstable performance in terms of the accuracy metric, with a standard deviation of 10.3%, 11.8%, and 8.0% for predicting the percentage of HighMI in Bipolar, Borderline and Control users. On the other hand, $\mathtt{C}$ shows a comparatively stable performance for the same numbers with only 0.4%, 0.7% and 0.3% standard deviation in accuracies respectively. I summarize the comparison values of the two classifiers in Table 5.13. Thus, while we do not see a drastic improvement in classification accuracies between $\mathtt{C}$ and $\mathtt{C}_0$, these results demonstrate the stability of the semi-supervised learning based classifier $\mathtt{C}$ especially in the face of limited availability of ground truth labeled data.

Examining Psycholinguistic Features

To understand the prominent psycholinguistic features of the C classifier, Table 5.14 summarizes the statistically significant features and their values for the two mood instability classes HighMI and LowMI. Broadly, I note that the mean occurrences of each of the psycholinguistic features is substantially higher in the timelines of users classified as HighMI as compared to those inferred to show LowMI.

To start with, we observe that the features under affective attributes, especially *anger, negative affect*, and *positive affect* show significant contribution towards the classification model. This agrees with the intuition that, individuals having traits of mood instability are likely to be expressive and use affective words. Looking at the class-wise differences in Figure 5.6 which plots the distribution of affective features across individuals. For positive affect, HighMI individuals show a substantial higher median than LowMI individuals (0.24 vs. 0.04). Likewise, a similar trend is observed for negative affect (0.20 for HighMI and 0.03 for LowMI individuals).

Returning to other psycholinguistic features described in Table 5.14, we find that cognitive attributes like *negation*, *discrepancies*, *cognitive mechanics*, *certainty* and *tentativeness* stand out, distinguishing the two mood instability classes. First, we observe that HighMI users show greater usage of cognitive attributes and perception. This finding aligns with prior work, which associates higher use of cognitive and perceptive words with emotional upheavals, and self disclosure about psychological conditions [475]. Next, the HighMI users show heightened self-attentional focus as illustrated in the usage of 1st person singular pronoun features; this value is significantly lower in the case of the users classified to show LowMI. Self pre-occupation is observed in individuals challenged with many mental health concerns, who in turn, in many cases, may also exhibit high instability in their emotional states [126]. In terms of temporal references, the HighMI users show a greater focus on here and now, indicated in the high usage of present tense words. Further, the occurrences of lexical density features such as *verbs* and *adverbs* in HighMI is almost 600% as compared in

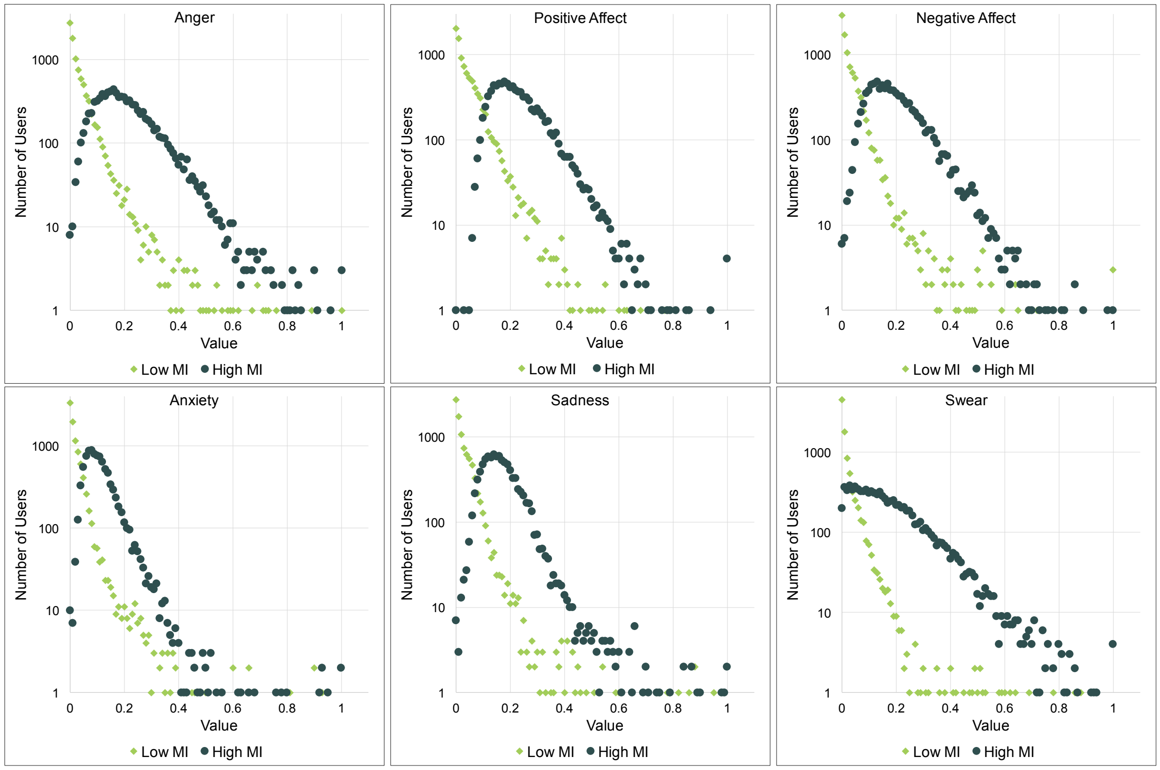


Figure 5.6: A comparative representation of the distribution of the values of LIWC affective attributes in users classified to be of HighMI or LowMI.

LowMI, indicating that individuals with higher mood instability tend to express themselves via more complex narratives, as also known from prior work in psycholinguistics [126].

Analyzing Mood Instability on Twitter

The final set of results include an analysis of the linguistic markers of mood instability as manifested in the target datasets from Twitter. Table 5.15 reports the top occurring, most relevant n-grams (n=2) based on their Log Likelihood Ratio (LLR) across two classes. In doing so, I also investigate whether and how the usage of different n-grams vary across the classes of mood instability. We observe that certain n-grams reported here, agree with the distribution of the psycholinguistic features that are significantly distinct across the two classes. For instance, ‘argue’ which occurs predominantly in HighMI, is categorized under the *social* and *anger* features in the psycholinguistic lexicon LIWC. In fact, these specific features of *anger* and *social* occur almost 6 times more frequently in HighMI as compared

Table 5.14: Psycholinguistic categories and their distribution across the two classes of mood instability. Only significant features for classifier C are reported here, with their score. Statistical significance is reported after Bonferroni correction ($\alpha = .05/50$).

Category	High MI		Low MI		p-value	Score
	Mean	Stdev.	Mean	Stdev.		
Affective Attributes						
Anger	0.21	0.12	0.04	0.05	***	33.37
Negative Affect	0.20	0.10	0.03	0.04	***	31.51
Positive Affect	0.24	0.10	0.04	0.05	***	35.46
Sadness	0.17	0.08	0.03	0.05	***	13.57
Swear	0.17	0.14	0.02	0.05	***	13.57
Cognitive Attributes						
Causation	0.18	0.08	0.03	0.04	***	26.38
Certainty	0.34	0.12	0.05	0.06	***	50.81
Cognitive Mech	0.39	0.12	0.07	0.07	***	60.86
Discrepancies	0.30	0.11	0.05	0.05	***	47.91
Negation	0.33	0.12	0.05	0.06	***	54.91
Tentativeness	0.20	0.09	0.03	0.04	***	31.41
Perception						
Feel	0.17	0.08	0.03	0.04	***	24.32
Insight	0.16	0.07	0.03	0.04	***	21.62
Percept	0.24	0.13	0.04	0.05	***	40.48
See	0.13	0.07	0.02	0.03	***	20.87
Interpersonal Focus						
1st P. Singular	0.28	0.12	0.05	0.06	***	46.90
2nd PP.	0.20	0.11	0.04	0.05	***	26.69
3rd PP.	0.10	0.07	0.01	0.03	***	15.70
Indefinite P.	0.36	0.12	0.06	0.06	***	59.35
Temporal References						
Past Tense	0.23	0.11	0.03	0.04	***	39.71
Present Tense	0.41	0.12	0.07	0.07	***	69.02
Lexical Density and Awareness						
Adverbs	0.39	0.13	0.06	0.07	***	65.03
Verbs	0.43	0.12	0.07	0.07	***	70.77
Exclusive	0.33	0.12	0.05	0.05	***	57.33
Inclusive	0.25	0.10	0.05	0.05	***	37.68
Preposition	0.37	0.13	0.07	0.07	***	56.78
Social/Personal Concerns						
Bio	0.16	0.07	0.03	0.04	***	21.52
Body	0.17	0.08	0.03	0.04	***	25.97
Death	0.09	0.07	0.02	0.04	***	12.98
Humans	0.13	0.07	0.02	0.04	***	17.12
Sexual	0.12	0.11	0.02	0.04	***	19.07
Social	0.300	0.11	0.05	0.06	***	46.18

to LowMI. This indicates, users in HighMItend to be more expressive and argumentative on social media, such as "*We curse, fight, kiss, hug, We text, talk, argue, laugh, We smile, We love. That's just us!*". Similar is the case with 'afraid', which exists under the *anxiety* feature in LIWC. We observe that a few phrases related to pregnancy and child birth, such as 'baby born', 'birth', 'feeding' and 'pregnant' are predominant in HighMI. Some example tweets include, "*Everyones pregnant or married and I'm..*", "*looking at my stomach, I can't believe I'm pregnant and I'm really having my own baby*", "*hungry + sleepy is a bad combination*

Table 5.15: Log Likelihood Ratios (LLRs) of top 20 n-grams more frequent in the posts of users classified as HighMI (left), LowMI (center) and comparably frequently in both (right).

n-gram (High>Low)	LLR	n-gram (Low<High)	LLR	n-gram (High=Low)	LLR
argue	1	finance	-1	followed people	0.03
awww	1	fountain pen	-1	red	0.04
baby born	1	global investment	-1	chocolate	0.04
birth	1	gus music	-1	smile	0.04
eyebrows	1	profit	-1	healthy	0.04
failed	1	health equity	-1	unfollowed automatically	0.04
fall asleep	1	health money	-1	followed	0.04
feeding	1	health money bitcoins	-1	goodbye	0.04
funding	1	investment plan	-1	learning	0.05
hip hop	1	irregular traffic	-1	adorable	0.05
hurting	1	management	-1	birthday	0.05
playing woman	1	millionaire	-1	relationships	0.06
pregnant	1	pension	-1	challenge	0.06
pressure	1	perfect money	-1	holidays	0.06
racism	1	remixes	-1	goodnight	0.06
republicans	1	single mother	-1	magic	0.03
suicide	1	fastest investments	-1	creative	0.02
fucked	0.41	equity careers	-1	lips	0.02
racist	0.49	entertainer	-1	thankful	0.01
favorite	0.47	download new	-1	thanks	0.01

for a pregnant woman." and "*this baby hurting my damn back. im not having any more kids*". This concurs with prior literature, on the association of mood instability with pregnancy related conditions [383]. We also observe the presence *swear* words, like 'fucked' and words associated with highly negative depressive acts and forms of expression as well as low self-esteem in HighMI, like 'suicide', 'hurting' and 'failed'. Example tweets here include, "*I can't believe my suicide is delayed*" and "*everybody likes hurting me all the time*"

On the other hand, the top n-grams from LowMI, contain some health, career and money related phrases like, 'health money', 'finance', 'perfect money', 'entertainer', 'equity careers', 'millionaire', 'pension'. This may indicate a tendency of LowMI Twitter users to engage in discussing more general life and lifestyle oriented topics, such as in tweets like, "*what would you buy if you became a multi-millionaire overnight*". In other words the lower presence of these n-grams in HighMI may indicate a relatively greater detachment of these users from the day-to-day realm. In contrast to pregnancy related words in HighMI, we find, 'single mother' occurs as a top n-gram in LowMI. Such a contrast interested us, and as I drill down to the corresponding tweets, we find some expression of dis-inhibiting opinions and disclosures relating to people's personal lives, such as "*my single mother* *worked without a penny from 'him' because she changed her lifestyle to be there for me. the real sign of a mother.*", "*of course a woman who was a poor single mother until she worked her way out of poverty can't possibly comment.*", "*being a single mother works with her being a really independent fierce woman who didn't give up her motherhood for her*" and campaign oriented tweets such as, "*retirement income plan for single mother*". Among the n-grams which occur almost equally in both the classes, we find quite a few phrases related to greetings, vacation and occasion, for example: 'goodbye', 'thanks', 'birthday', 'goodnight', 'holiday' etc. which are typically expected to surface in many casual social media chatter.

5.1.4 Discussion

This study presented a novel machine learning approach for inferring psychological states (mood instability) of an individual based on their social media data, leveraging dense, high fidelity ground truth information from an independently acquired active sensor—specifically ecological momentary assessments (EMAs). I demonstrated passively gathered social media data can be utilized to build an enriched and scalable mood instability classifier.

The results show that the proposed semi-supervised learning approach makes significant contributions towards exploring how very small samples of actively sensed data can be augmented with large-scale social media data to detect individuals' binary mood instability status (low, high), robustly, with 96% accuracy and F-1 score. The proposed semi-supervised learning method can detect high mood instability, that, in comparison to a suitable control population, reveals meaningful linguistic 'signatures' in the social expression of Twitter users who self-disclose to suffer from bipolar or borderline personality disorder. The method indicates that the bipolar and borderline personality disorder populations exhibit high mood instability, with almost twice the likelihood of a control population; an observation aligning with relevant literature in psychology [22, 58, 141].

This study highlights an unconventional, yet creative mechanism to rethink certain study designs within the ubiquitous computing community. Many of these studies typically employ sophisticated and highly engineered systems for sensing behaviors, moods, and activities of individuals. Incentives are also needed to be built into the design to maintain participant compliance while reducing burden. I show that with access to voluminous naturalistic social and behavioral data gathered from social media unobtrusively and passively, these study designs may be revisited. Existing sensing frameworks that employ small-scale active data collection could thereby tackle the challenges of scalability to large populations and to extended periods of time, by utilizing complementary social media data of the population being studied. Moreover, unlike most active sensing paradigms, I demonstrate that with social media data, we can leverage access to the rich context within which activities and moods unfold and are expressed, such as their social and behavioral underpinnings. Such information can be immensely helpful in many health sensing applications [3], beyond the investigations presented in this paper.

By borrowing a semi-supervised learning approach from machine learning, this study builds on the success of these methods [684]. While fully supervised approaches (e.g., regression and classification) are routinely used for health sensing [187, 652], this work reveals that a semi-supervised approach can promisingly tackle the challenges around paucity of labeled data (e.g., individuals suffering from a health condition), by incorporating easily accessible unlabeled examples. Moreover, the findings suggest that a semi-supervised classification approach improves the performance of a seed fully supervised classifier, both in terms of robustness and confidence, indicating the applicability of the proposed approach in real-world affect and mood inference tasks, beyond laboratory studies.

This methodology can be applied in a variety of other health sensing problems, especially problems challenged by limited access to groundtruth. More generally, over half of American smartphone users are reported to spend an average of 144 minutes per day browsing their mobile devices, aiming to stay socially connected with their friends [262]. These users often identify as quantified selfers, which includes tracking signals from a range of wearable sensors (such as heart rate, body acceleration or physical location). Given the popularity of social media technologies, this study shows a mechanism to bridge the gap between individuals' online representation and actual physical and emotional status, and how they can mutually benefit each other in health status sensing tasks.

Broadly, this study advances the vision proposed by Estrin [207] and Zhang et al. [681] around developing approaches within the precision medicine context, that can integrate multiple forms of technology facilitated sensed data into improved understanding of health and wellbeing. The integration of EMA and social media technologies will enable us to better understand the early signs that may indicate forthcoming risk to unusual shifts in mood or another adverse health episode. To realize this goal, software infrastructures to enable automated social media sensing of health, alongside other forms of sensing may be developed, akin to the Aware [212] and SenSocial [412] frameworks that allows unobtrusive logging of passive data centered around people's smartphone activity. This study reveals the potential of novel systems and interventions that can proactively monitor wellbeing may be designed and deployed. These can be in the form of self-tracking tools that allow self-reflection for individuals, or in the form of interfaces that could be used by clinicians and caregivers so as to direct timely and personalized help [676].

The findings reveal that social media can function as a source of passively and unobtrusively sensed data to infer mood instability in individuals, and can significantly augment existing small-scale active sensing techniques. The implications of this study are situated within precision medicine, around how multisensor integration of signals relating to health can improve the assessments and understanding of challenging mental health concerns. Given the widespread adoption of social media technologies, this study bridges the gap in observations between individuals' online representation and actual physical and emotional status, and how they can mutually benefit each other in health status sensing tasks.

5.2 Leveraging Multimodal Sensing Data to Impute Missing Social Media Data

Understanding *why* and *how* individuals feel, think, and act is a key topic of interest among researchers from a variety of academic disciplines, such as psychiatry, psychology, sociology, economics, and anthropology [378]. Typically, studies of human behavior have largely relied on self-reported survey data. However, these approaches bear limitations, for example, survey data suffers from subjective assessments, recall and hindsight biases. Active and passive sensing technologies overcome these challenges by recording psychological states and behavior in-the-moment [96]. However, sensing-based approaches require diverse, extensive, and rich data via multiple modalities to obtain comprehensive information about an individual's state and context [96]. In comparison to active sensing such as Ecological Momentary Assessments (EMAs), passive sensing techniques mitigate the challenges of compliance and response burden. However, passive sensing paradigms are typically limited to capturing behaviral data only during the study or active data collection period [565]. Such a drawback could be overcome by leveraging social media data, which is an inexpensive, unobtrusive, and naturalistic means to gather both present and historical data of individuals [529]. Additionally, social media data is a form of *verbal sensor* to capture people's linguistic expressions, therefore, a complementary means to infer psychological dynamics of individuals [166, 528, 564].

That said, the availability and quality of social media data widely vary on people's social media use. Passive consumption is often more prevalent than active engagement, leading to sparsity in data over extended periods of time. Consequently, studies either focus on a very active participant cohort — hurting *generalizability* and *recruitment*, and introducing *compliance bias*, or disregard those with no or only limited social media data — hurting *scalability*. Additionally, everybody is not on social media, and its use is typically skewed towards young adults [479]. Yet, research may require to study demographics where social media is less prevalent. Again, gathering social media data also presents engineering challenges due to platform restrictions.

Therefore, this study aims to address the challenges of missing sensing streams (here, social media) in multimodal sensing studies of human behaviors. This study is theoretically grounded in the Social Ecological Model [102] that posits human behaviors have social underpinnings, and are deeply embedded in the complex interplay between an individual, their relationships, communities, and societies.

This study examines: *How to leverage the potential of social media data in multimodal sensing studies of human behavior, while mitigating the limitations of acquiring this unique data stream?* I address this question within the Tesserae project [406], a multisensor study that aims to predict psychological constructs using longitudinal passive sensing data of 757 information workers.

Focusing on the participants whose social media data is not available, this study proposes a statistical framework to model the latent dimensions which could have otherwise been derived, had their social media data stream been available. Specifically, I impute *missing social media features* by learning observed behaviors from other passive sensor streams (bluetooth beacons, wearable, and smartphone sensors). I employ a range of state-of-the-art machine learning models, including linear regressions, ensemble tree-based regression, and deep neural network based regression. After demonstrating that the imputed social media features closely follow actual social media features of participants (average correlation of 0.78), I evaluate the efficacy of the social media imputation framework. I compare pairs of statistical models that predict a range of common (or benchmark) individual difference variables (psychological constructs like personality, affect, and anxiety) — one set of models being those that use imputed social media features alongside other passive sensor features, and the other set that does not use these imputed signals. The findings suggest that the imputed social media features significantly improve the predictions by 17%.

Summarily, this study shows that the proposed framework can augment the range of social-ecological signals available in large-scale multimodal sensing studies, by imputing latent behavioral dimensions, when one sensor stream (that is, social media data stream) is *entirely unavailable* for certain participants. I discuss the implications of this study as a methodological contribution in multimodal sensing studies of human behavior, within the sensing research community.

5.2.1 Study and Data

Our dataset comes from the Tesserae project that recruited 757 participants (section 4.1). Note that this study was conducted while the participation was ongoing, so it uses data till August 21, 2018. Randomly Randomly selected 154 participants were “blinded at source” whose data was put aside only for external validation at the end of the study. This study only concerns the data of the remaining 603 “non-blinded” participants in the study.

The dataset consists of 350 males and 253 females, where the average age is 34 years (stdev. = 9.34). In education, the majority of the participants belong to have college (52%) and master’s degree (35%) education level.

Passively Sensed Data. The participants were enrolled over 6 months (February to July 2018) in a staggered fashion, averaging at 111 days of study per participant. Table 5.16 reports the descriptive statistics of the number of days of passively sensed data that was collected per participant through each of the sensor streams. Per participant, there is an average of 42 days data through bluetooth beacons, 108 days data through wearable, and 101 days of data through a phone application.

Out of the 603 non-blinded participants, 475 authorized their Facebook data. This data can be broadly categorized in two types—ones that were self-composed (e.g., writing a status update or checking into a certain location), and ones that they received on shared updates on their timeline. Comprehensively, Facebook data consists of the updates on participants’ timelines, including textual posts, Facebook apps usage, check-ins at locations, media updates, and the share of others’ posts. The likes and comments received on these updates

Table 5.16: Descriptive statistics of # days data collected.

Type	Range	Mdn.	Std.
Study Period	16:205	99	46.7
Bluetooth	1:159	37	32.6
Wearable	5:206	94	46.9
Smartphone	1:206	93	52.4
Social Media	110:4756	2923	1474

Table 5.17: Descriptive statistics of the Facebook dataset.

Type	Mdn.	Std.
Likes Rcvd.	1,139	5,277.85
Comms. Rcvd.	316	1,383.69
Self-posts	137	511.80
Self-comments	55	334.16
Self-Words	2,374	13,718.56

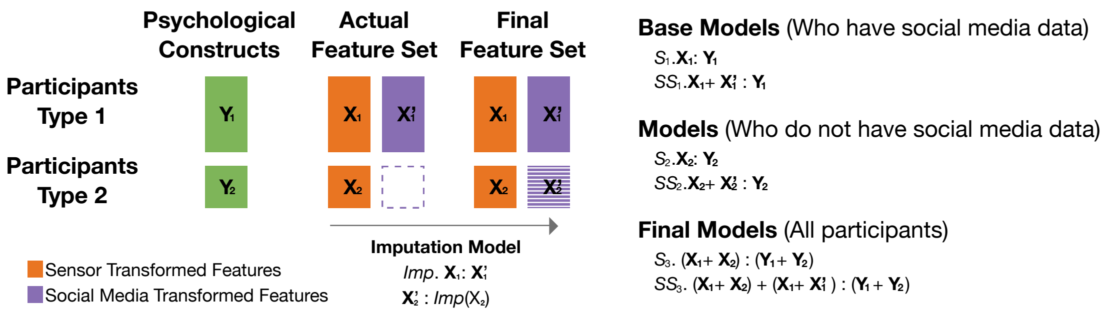


Figure 5.7: A schematic overview of the statistical models built to evaluate the effectiveness of imputation.

on the participants' timelines were also collected. Table 5.17 summarizes the descriptive statistics of the Facebook dataset. Temporally, the data dates back to October 2005, and the number of days of data per participant averages at 1,898 days — giving us a sense of the historical data that Facebook allows us to capture.

5.2.2 Methods

Feature Learning Framework

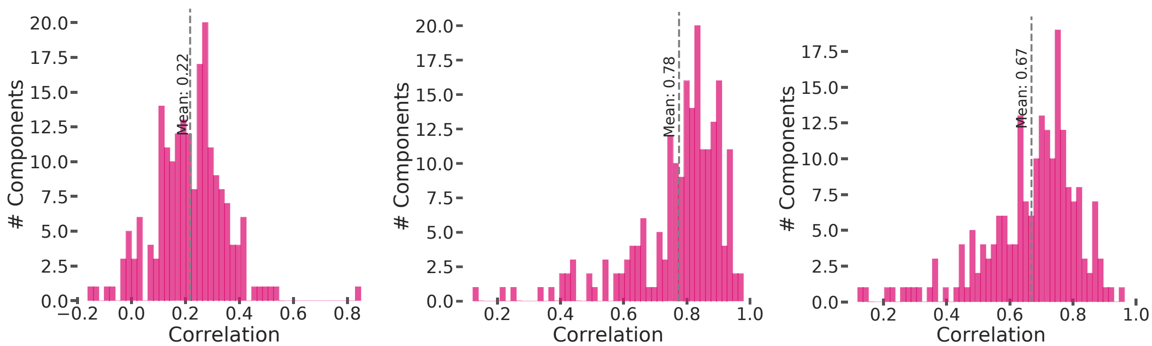
I build a feature learning framework to address the challenge of missing social media data for 128 participants. Figure 5.7 shows a schematic overview of the prediction models of psychological constructs that are used to evaluate the effectiveness of the imputing missing social media transformed features. I briefly mention the three algorithms that are consistently used throughout the study.

Linear Regression (LR) Linear regression adopts a linear approach to model the relationship between the independent and dependent variables [566]. Specifically, wherever applicable, I employ linear regression with L1/L2 regularization to prevent overfitting and to avoid bias introduced due to the inter-dependence of independent variables [687].

Gradient Boosted Regression (GBR) Gradient boost technique conducts regression in the form of an ensemble of weak prediction models, which are typically decision trees [199, 426]. It optimizes the cost function by iteratively choosing a function that points in the negative gradient direction. This study uses gradient boost on an ensemble of decision tree regressors, by varying the number of decision trees between 100 and 1000, with each tree of maximum depth as 3.

Multilayer Perceptron Regression (MLP) Neural network regression suits in problems where a more conventional regression model cannot fit a solution. I use the multi-layered perceptron (MLP) technique that works in a feed-forward fashion (no cycles) with multiple internal layers [522]. The model learns through a method called backpropagation [368], and follows a fully connected (dense) deep neural network architecture. Wherever applicable, I use two internal layers and tune the number of nodes in them between 36 and 216 for the neural network regression models.

The above three algorithm choices are motivated by the fact that they essentially cover a broad spectrum of algorithm families spread across linear regression, non-linear regression, decision trees, ensemble learning, neural networks, and deep learning. I quantify the prediction accuracy of psychological constructs as the Symmetric Mean Absolute Percentage Error (SMAPE), which is computed as mean percentage relative difference between predicted and actual values, over an average of the two values [316]. SMAPE values range between 0% and 100%, and lower values of error indicate better predictive ability. To obtain these, I first divide their datasets into five equal segments, and then iteratively train models on four of the segments to predict on the held-out fifth segment. I average the testing accuracy metrics to obtain the pooled accuracy metrics for the above algorithms. I refer to this technique as *pooled accuracy technique* and the corresponding outcomes as *pooled accuracy or error measures*. Within the training segments, I tune the hyper-parameters using a k-fold



(a) LR (b) GBR (c) MLP

Figure 5.8: Correlation distribution between PCA Components and Predicted PCA Components of Facebook

cross-validation ($k = 5$) technique.

Baseline Prediction with Passively Sensed Data. I first seek to establish if the presence of social media features improves prediction accuracy. On the same set of 475 participants who have social media data, I compare two models of predicting psychological constructs — 1) $\mathtt{S}_1$ uses 30 sensor features, and 2) $\mathtt{SS}_1$ combines 30 sensor features and 200 social media features. Table 5.18 reports the relative decrease in error for $\mathtt{SS}_1$ compared to $\mathtt{S}_1$. The relative decrease in error averages at 21% for LR, 26% for GBR, and 21% for MLP. In sum, adding social media features improves the predictions by an average of 22.4% across all the models and the psychological constructs.

Imputing Missing Social Media Features

The baseline prediction suggests that adding social media features *indeed* improves the prediction task of the psychological constructs. However, about one-quarter of the participants do not have social media data (see section subsection 5.2.1). This restricts us from leveraging a rich feature stream to predict such attributes for these individuals. To overcome this constraint, this study aims at *learning* certain latent behaviors that I could have otherwise inferred if I had access to their social media data.

I impute the social media features using the sensor features. For this, I build learning models on the sensor stream of the social media participants to predict their latent social media dimensions. That is, for every 200 social media feature, I build a separate model that uses the sensor features as the independent variables to predict the social media feature. I adopt k-fold cross-validation based hyper-parameter tuning. I use LR, GBR, and MLP to find the best algorithmic model, and quantify the pooled accuracy of the prediction models in terms of Pearson's correlation (r) between actual and predicted social media features. Levene's test between all the actual and predicted features reveals homogeneity of variance in the feature set [443]. This statistically indicates that the imputed social media transformed features are not arbitrarily generated.

Figure 5.8 plots the distribution of the pooled Pearson's correlation (r) between the actual and predicted values of social media transformed features. I find that the mean correlation across the components is 0.22 in LR, 0.78 in GBR, and 0.67 in MLP. All of these correlation measures are statistically significant at $p<0.05$. Comparing across the algorithms, GBR performs the best in predicting the latent social media dimensions. For the rest of the analyses, I use the GBR algorithm to impute the social media transformed features.

5.2.3 Results

Evaluating the Effectiveness of Imputation

Among 128 participants without social media data, I compare two prediction models of psychological constructs— 1) $\texttt{S}_2$ uses only sensor features of these participants, and 2) $\texttt{SS}_2$ combines sensor features and imputed social media features (as obtained above).

I compare the accuracy metrics of $\texttt{S}_2$ and $\texttt{SS}_2$ to deduce if imputing the social media features improves the task of predicting psychological constructs. Table 5.18 compares the prediction errors (SMAPE) for the three algorithms that I run in each of the models $\texttt{S}_2$ and $\texttt{SS}_2$. We find that for LR, the relative decrease in the error ranges between 6% (for openness) and 17% (for positive affect), averaging at 11%; for GBR, the relative decrease in the error ranges between 16% (anxiety) and 20% (extraversion), averaging 17%; and for MLP, the relative decrease in the error ranges between 6% (extraversion) and 21% (anxiety).

Table 5.18: Relative % decrease in SMAPE in prediction models using both sensor & social media features from ones using only sensor features. Positive values mean better prediction in $\mathtt{SS_n}$ than $\mathtt{S_n}$.

	$\mathtt{SS_1}$-$\mathtt{S_1}$			$\mathtt{SS_2}$-$\mathtt{S_2}$			$\mathtt{SS_3}$-$\mathtt{S_3}$		
Psy. Construct	**LR**	**GBR**	**MLP**	**LR**	**GBR**	**MLP**	**LR**	**GBR**	**MLP**
Personality Traits (BFI-2)									
Extraversion	10.6	28.4	16.6	8.4	20.1	6.4	12.8	19.5	3.6
Agreeableness	8.3	27.5	30.4	5.9	17.9	17.2	3.2	14.4	20.2
Conscientiousness	11.8	26.0	28.2	9.4	17.4	13.5	15.0	21.2	12.1
Neuroticisim	11.2	24.9	17.6	7.6	16.9	13.4	6.0	17.5	-13
Openness	10.0	25.1	33.8	6.1	15.6	16.9	5.4	15.3	3.1
Affective Measures									
Pos. Affect	33.8	26.2	8.06	16.6	18.1	18.4	8.6	14.5	21.5
Neg. Affect	38.8	24.7	24.04	16.1	15.7	9.7	8.4	11.8	16.4
Anxiety (STAI)	39.4	24.3	7.5	14.1	15.7	20.8	6.4	16.8	34.4
Mean	20.5	25.9	20.8	10.5	17.2	14.5	8.2	16.4	12.3

Therefore, the imputed social media features improved the prediction by an average of 14% across all models and measures.

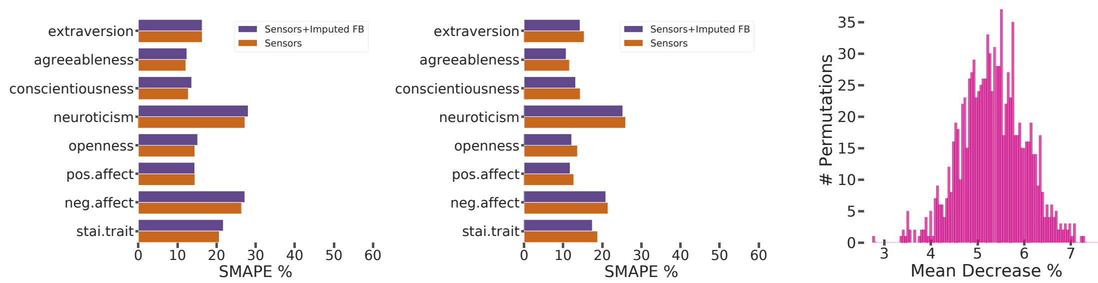
Finally, on the entire dataset, I build two *Final Models* to evaluate the overarching effectiveness of imputation— 1) $\mathtt{S_3}$ uses sensor features of all participants, 2) $\mathtt{SS_3}$ uses Facebook features of all participants. In this model, for those who have Facebook data, I use Facebook features, and for the rest, I use imputed Facebook features.

I compare the prediction accuracy of the $\mathtt{SS_3}$ and $\mathtt{S_3}$— this gives us an estimate of how this sort of imputation framework influences the overarching task of predicting psychological constructs in multimodal studies (see Table 5.18). We find an average improvement in prediction by 8.2% in LR, 16.4% in GBR, and 12.3% in MLP.

Hypothesis Tests for Robustness

After evaluating the imputation models, I measure its robustness. I compare the effectiveness of the imputed sensing stream against two other imputation approaches applied to those 128 participants without social media data.

Mean Imputation. This approach imputes social media features as the mean value of the corresponding feature sets. I build prediction models of psychological constructs as described in the previous subsections. This method draws on prior studies which adopted similar



(a) Mean Imp. (b) Random Imp. (c) Permutation Test

Figure 5.9: (a&b) SMAPE of prediction models with sensor features ($\mathtt{S}_3$) vs. those with sensor and (a) mean- and (b) random- imputed features, (c) Reduction in SMAPE in several permutations of randomly imputed social media features compared to $\mathtt{S}_3$.

approaches of imputing missing features using static measures of central tendencies, such as mean or median of the feature sets [184].

Randomized Imputation. This approach imputes the social media transformed features as random values from the corresponding feature sets. I repeat such a randomization for a 1000 times, and in each case compare the prediction with *Final Model* $\mathtt{S}_3$. This method emulates a permutation test [17], and checks for robustness of the imputation effectiveness, by testing the null hypothesis that randomly imputed sensor streams are better than that imputed by the statistical framework.

Figure 5.9 shows the SMAPE of these models compared to $\mathtt{S}_3$. While the imputation shows an average improvement in SMAPE by 16% on the *Final Model* ($\mathtt{S}_3$) (see Table 5.18), the same improvement for *Mean Imputation*-based model is -3.10% and *Randomized Imputation*-based model is 5.34%, suggesting minimal (or no) improvement in these two models. Permuting on the randomized imputations a thousand times, we observe that in terms of prediction error, the imputations are *never* outperformed by the randomized imputations in those thousand permutations. Essentially, this *rejects* the null hypothesis that *our imputation is only more effective than randomly generated imputations by chance*.

In conclusion, the findings reveal that passively sensed multimodal data streams can be used to not only impute latent social media dimensions, but also to augment these latent features in building better prediction models that infer psychological constructs.

We consistently observe similar trends in the improvement of prediction accuracies by integrating the social media features (both actual and imputed) with the sensor-transformed features.

5.2.4 Discussion

Theoretical and Practical Implications. This study proposes an analytical framework of imputing a missing sensing stream (here social media) in multimodal sensing studies. We evaluate the effectiveness of this imputation by predicting psychological constructs through a variety of state-of-the-art algorithms. At a higher level, the imputation framework is grounded on the Social Ecological Model that construes interdependence among individuals, their behavior and their surroundings and environment [548, 647]. This implies its applicability not only in theory but also in practice (context and activity as captured and observed through passive sensing modalities). Our findings reveal the robustness of imputation by comparing with permutation tests and random- and mean- imputation. Such a framework can potentially be used in studies where there is similar theoretical grounding (around a focus on comprehensive social ecological signals), and an opportunity to infer psychological attributes.

The findings suggest that integrating social media features improves the prediction of psychological constructs. This aligns with prior work on the potential of social media (both individually as well as in tandem with other passive sensors) in predicting these measures [166, 529, 564]. However, social media data may not be available for the participants. the proposed imputation method addresses this gap by computing latent social media dimensions, which can be used to improve such machine learning-based prediction tasks of human behavior.

Following this framework, existing datasets that include multimodal sensing, but do not have social media streams for some participants, can now be retrained for better predictions. While this study only focuses on predicting psychological constructs, the same method can be extrapolated to predict other measures of human behavior as well. Not being limited to a single algorithm, the framework shows the consistency in the findings across a variety of algorithm families. It is not constrained by the choice of machine learning algorithms, which typically vary depending on the characteristics of the dataset and the distribution of the individual difference variables.

Additional sensing streams and features can be plugged into the framework. However, it remains interesting to study whether the additional sensors improve the imputation models. For instance, sensing technologies that capture conversations [499] among individuals in social settings would plausibly improve predicting latent social media features, on the rationale that it captures another set of dimensions in the social ecological framework — offline social interactions.

Ethical Implications. This study cautions against its misuse as a methodology to surveil or infer individual behaviors. This study intends to model latent dimensions that can assist prediction tasks in multimodal sensing studies, by being internal to the pipeline of the prediction system. However, these latent dimensions do not necessarily translate to or are indicative of actual individual behaviors on social media, and therefore such inferences cannot be drawn from the imputed social media features about the individuals.

This study does not unpack why some participants did not share social media data. It could be because they do not use social media, or because they do not intend to share for privacy reasons. Whether social media features should be imputed for these individuals can constitute a debated topic. This is because such an imputation approach, when applied to make predictions of sensitive individual difference variables and incorporated into larger systems (e.g., targeted advertising), can be perceived as a violation of the very privacy considerations that spurred them to not share social media data in the first place.

CHAPTER 6

INTROSPECTING INTO ONLINE-DATA-DRIVEN OFFLINE INFERENCES

So far, I highlighted the potentials of social media in inferring wellbeing. However, we need to recognize that these computational assessments have several potential real-world consequences. For example, these assessments may be instrumented in taking high-risk decisions, such as hiring or firing employees at workplaces. Therefore, we need to be careful about these assessments, and meaningfully understand what we are measuring. In this regard, this chapter aims to investigate and interpret the online-data driven offline metrics.

The first study examines how we can conduct person-centered predictions using social media data by leveraging multimodal sensing to contextualize the offline contexts of individuals. This paper aims to balance the trade-off between one-for-each and one-for-all models by clustering individuals on mutable behaviors and conducting cluster-specific predictions of psychological constructs. This work begins with a hypothesis that complementing social media with data with offline sensor data can help to personalize and improve predictions. However, the findings reveal mixed observations with respect to significant improvement in certain psychological constructs (e.g., sleep quality) and no improvement in others (e.g., cognitive ability). This study reveals the importance of taking a critical stance on evaluating the effectiveness before investing efforts in personalization.

The second study examines the characteristics of and factors explaining life event disclosures on social media. In particular, as social media platforms continue to evolve as archival platforms where individuals disclose several aspects of their lives for support, solidarity, maintaining and gaining social capital, and meeting therapeutic needs. I study what life events are disclosed on the year-long Facebook data of individuals in comparison to their self-reported life events in the same period. This study contributes a codebook to identify life event disclosures and builds regression models on event-centric and individual-centric factors that are associated with life event disclosures. The findings reveal that while all life events may not be disclosed, online disclosures reflect complementary information to self-reports. This study bears practical and platform design implications in providing support and sensitivity to life events.

6.1 Contextualizing Person-Centered Predictions with Social Media

The past few years have increasingly seen several passive sensing approaches to improve our understanding of human behavior both longitudinally and scalably. Simultaneously, research has utilized ubiquitous social media platforms as a "passive sensor" [529] and an unobtrusive source of behavioral data, which is self-recorded and self-initiated by individuals in naturalistic settings. Because this data contains language and social interactions, it is a unique form of *verbal* and *social* sensor, unlike several physical sensing modalities. A large body of research reveals the potential of inferring psychological constructs with social media [166, 252, 564].

However, social media data may not include the valuable contextual information that drives posting behaviors. For instance, even within the same emotional state, Facebook posting varies across individuals [399]. This interaction of various factors underscores the idea of the Social Ecological Model [102] in which psychological constructs are embedded in a complex interplay between individual, social, and environmental factors. Posting (or not posting) can be dictated by external factors that vary for every person. Therefore, social media sensing is unique in its sensitivity to factors driving an individual's self-initiation, motivation, and presentation. This between-person variability in data may impact predictions of an individual's underlying psychology, routines, and other personal attributes. Incorporating additional offline context that captures factors affecting online behavior could boost the ability of social media to predict individual outcomes.

Personalizations, where models are tuned and optimized for each individual [524] can overcome between-subject variability. Indeed, personalized modeling methods are gaining attention in the disciplines of social science, psychology, and health, including precision medicine and digital phenotyping [294, 309, 364, 454, 467]. Person-centered methods can glean a more comprehensive understanding of an individual, and in some cases explain their outcomes better than variable-centered or generalized methods (i.e. focusing on global variables that are measured through the same means for everyone in a target population) [391, 527, 670]. A variety of personalized predictions have also been conducted in computing, particularly in content recommendation and data mining [576]. Drawing on such approaches, we can consider predictions by building individual-level models. However, such approaches would be impeded due to the temporal sparsity of social media data (because individuals post on social media only at discrete intervals). Alternatively, we can consider stratifying individuals on demographic attributes such as age, race, and gender. However, these attributes are not only privacy-intrusive but are also static and exclusionary. Demographic attribute-based stratified modeling has been identified by the Fairness, Accountability, Transparency literature to reinforce stereotypes and existing societal biases, and even exacerbate them [292, 314, 498]. Additionally, these approaches may not have sufficient data for a particular demographic or marginalized group. Using dynamic features shared more broadly can be a better alternative.

This study avoids demographic based and personalized modeling shortcomings by embracing multimodal sensing in capturing behavior and context, in the form of "small data" about a person [207]. I propose a person-centered approach that leverages passively collected dynamic attributes spanning phone use, physical activities, mobility, and work behaviors. Data is collected from Bluetooth beacons, smartphones, and wearables. We then obtain clusters (or groups) of individuals who demonstrate similar combinations of multidimensional offline behaviors. Clustering individuals is theoretically motivated in that people's offline behaviors drive online (or social media posting) behaviors and vice-versa. Clustering allows to balance within-individual heterogeneity and between-individual homogeneity. This approach is a middle-ground between "one-for-each" and "one-for-all" models, thereby aiming to balance the drawbacks of extremely personalized (one or few individuals per model) which may not generalize or consolidate findings across individuals, and extremely generalized models, which face difficulty in generating precise predictions per individual and can also be impeded by variability in individual data quality and completeness. This study draws motivation from clustering based approaches previously adopted in digital phenotyping research with electronic health records to identify comorbidity of symptoms [188] and to compare and summarize clinical models [248].

This study broadly hypothesizes that contextualizing on offline and naturalistic behaviors can provide a degree of personalization and improve predicting psychological constructs with social media. Combining multiple sensing modalities in this manner can allow us to leverage complementary strengths of different sensing techniques. Additionally, this approach can provide a theoretical lens of understanding the interaction between offline and online behaviors that is useful in both research and in practice. This study, therefore, targets the below *research aims*:

Aim 1: To predict psychological constructs with social media in a person-centered approach of contextualizing people's offline physical behaviors.

Aim 2: To evaluate and compare contextualized and generalized prediction models.

Aim 3: To examine how social media language associates with offline behaviors.

I use data from the Tesserae project [406], where 572 participants provided social media (Facebook) data. Consented participants provided self-reported measures of psychological constructs of *cognitive ability, personality traits, affect*, and *wellbeing*, which serve as ground-truth in this study. I use this data to achieve the *aims* above, through three-fold contributions:

First, this study contributes an approach of building *contextualized* person-centered models that predict psychological constructs from naturalistic passive data describing a multitude of contextual factors. I build *contextualized* models trained on each cluster's social media data and compare the performance against *generalized* models trained on the entire social media dataset of all participants, as is typically done.

Second, this study provides insights about the relative performance of predicting psychological constructs with generalized and contextualized models. The evaluations reveal that contextualized predictions show a significant increase in predicting anxiety, sleep, and personality traits, whereas no significant difference in predicting affect, and a significant decrease in predicting cognitive ability.

Third, I critically discuss the tradeoff between personalization and statistical power, and the importance of evaluating the costs and benefits of personalizations as implications in research and practice. This study construes that the utility of contextualizing on offline behavior for social media based predictions relies on the strength of the theoretical associations between a construct of interest and offline manifestations of the construct. Additionally, personalized models are not only costly but may also be impacted by the limitations associated with smaller training data sizes compared to generalized models. Theoretically, this work can be useful in behavioral modeling in emergent fields like human-centered machine learning, as well as to generate hypotheses for future investigations that leverage the relationship between passively sensed behavior and psychological constructs.

6.1.1 Study and Data

The data for this study comes from the Tesserae project (section 4.1) [406].

Self-Reported Data

The enrollment process consisted of an initial survey questionnaire related to demographics (age, gender, education, type of occupation, role in the company, and income), and survey questionnaires of self-reported psychological constructs of cognitive ability, personality traits, affect, anxiety, and sleep quality. Table 6.1 summarizes the distribution of the self-reported data within the dataset, where we find a reasonable distribution within demographics and psychological traits among the participants.

Figure 6.1 presents Pearson's r between the psychological constructs (Cognitive Ability, Personality Trait, and Affect and Wellbeing variables) and regression (R^2) results for the demographic and job-related characteristics as independent variables, and the psychological constructs as dependent variables. These correlations between psychological constructs, mirror prior literature. The observed positive correlation ($r=0.79$) between Neuroticism and Anxiety and Negative Affect is consistent with past research showing positive associations between elevated Neuroticism and mood and anxiety disorders [456]. The strong positive correlation ($r=0.67$) between Anxiety and Negative Affect is consistent with past research showing a strong co-morbidity between depressed mood and elevated anxiety [457]. Extraversion and Positive Affect ($r=0.54$) have also been found to be strongly positively correlated [385, 386]. Other past research has also found a negative association ($r=-0.51$) between Positive Affect and Anxiety [81], as well as a negative association ($r=-0.41$) between Conscientiousness and Anxiety [209]. All other inter-construct correlations are moderate at $|r|<0.40$. Next, looking at the association between demographic and job related variables and the psychological constructs, we observe only modest associations, with the strongest association being between Income bracket and the Shipley Crystallized Vocabulary scale ($R^2=0.05$). When all demographic and job-related variables are included in a regression model predicting the psychological constructs, we still observe only modest predictive performance for all psychological constructs (all $R^2<0.12$).

Passive Sensing Data for Offline/Physical Activity

This study collected offline/physical behaviors of individuals through three modalities of passive sensing, bluetooth beacon, wearable, and smartphone.

Table 6.1: Descriptive statistics of self-reported demographics and psychological constructs of participants.

Covariates	Value Type	Values / Distribution
Demographic Characteristics		
Gender	Categorical	Male | Female
Age	Continuous	Range (20:68), Mean = 34.90, Std. = 9.74
Education Level	Ordinal	5 values [HS., College, Grad., Master's, Doctoral]
Job-Related Characteristics		
Income	Ordinal	7 values [<$25K, $25-50K, ... , >150K]
Tenure	Ordinal	10 values [<1 Y, 1Y, 2Y, ... 8Y, >8Y]
Supervisory Role	Boolean	Non-Supervisor | Supervisor
Cognitive Ability (Shipley scale)		
Fluid (Abstraction)	Continuous	Range (0:24), Mean = 16.98, Std. = 2.84
Crystallized (Vocabulary)	Continuous	Range (0:40), Mean = 33.15, Std. = 4.11
Personality Trait (BFI scale)		
Openness	Continuous	Range (1.17:5), Mean = 3.82, Std. = 0.61
Conscientiousness	Continuous	Range (1.42:5), Mean = 3.88, Std. = 0.66
Extraversion	Continuous	Range (1.58:5), Mean = 3.42, Std. = 0.69
Agreeableness	Continuous	Range (2.08:5), Mean = 3.89, Std. = 0.56
Neuroticism	Continuous	Range (1:4.92), Mean = 2.46, Std. = 0.79
Affect and Wellbeing		
Pos. Affect	Continuous	Range (13:50), Mean = 34.53, Std. = 6.05
Neg. Affect	Continuous	Range (10:43), Mean = 17.52, Std. = 5.35
Anxiety	Continuous	Range (20:72), Mean = 38.13, Std. = 9.49
Sleep Quality	Continuous	Range (0:19), Mean = 6.65, Std. = 2.59

Social Media Data

Among the social media data streams, Facebook is the most popular social media platform [262], and that its longitudinal nature has facilitated several social media studies of understanding individual differences [164, 399, 529], it suits the particular problem setting. Facebook is also the most prevalent social media stream in the dataset, with 572 participants authenticating their Facebook data, among which 32 participants have no entries in their Facebook data — this study uses the remaining 540 participants' Facebook data for measuring psychological constructs.

6.1.2 Feature Engineering

I derive machine-usable features from the raw multimodal sensing data. I draw on prior work to derive features that have shown theoretical relevance in measuring psychological constructs [651, 653]. This section explains the features: first, those derived from physical

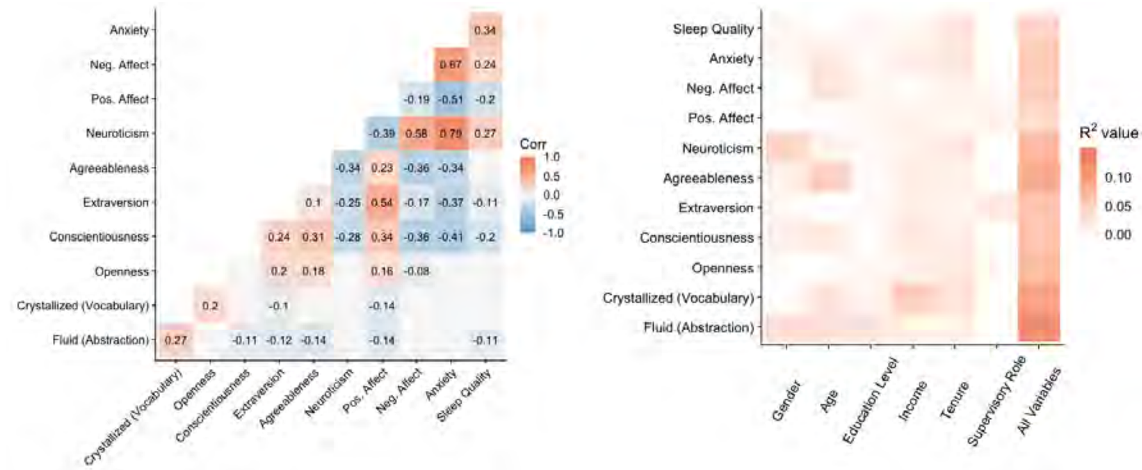


Figure 6.1: (Left): Pearson's r between psychological constructs, non-significant correlations (p>0.05) are left as blank.
(Right): Regression (R^2) results for the demographic and job-related variables as independent variables, and the psychological constructs as dependent variables (right sub-figure). The "All Variables" column provides regression results for the psychological constructs with all demographic and job-related variables included in the model.

sensors, followed by those derived from social media.Out of the 757 participants' data, I set aside a random sample of 6.7% (50) participants' data as the held-out dataset for final evaluation purposes. I conduct feature engineering, and build (train and validate) the models within the remaining 93.3% (704) participants' data.

Deriving Features from Physical Sensor Dataset

From the passive sensor data streams, I derive a variety of features that are related to participants' activity, sleep, and other physical behaviors, as summarized below:

Step Count The Garmin wearable collects fitness-related measures such as the daily step count of participants [234].

Physical Activity The smartphone app installed on participant smartphones used the Google Activity Recognition API [255] to identify physical activity at regular intervals. For each individual, I obtain durations of (1) *high* and (2) *moderate* intensity activities using the Metabolic Equivalent of Task metric from the wearable data [637].

Mobility The smartphone application continually recorded the GPS coordinates of the individuals. I derive the number of locations and the distance traveled between each location based on a 15-minute pooling window. I use this data to also derive (1) information on the total distance traveled each day, (2) the number of distinct locations visited, and the (3) maximum and (4) average distance from home traveled by an individual each day.

Phone Use Activity The installed smartphone app recorded the activity of smartphone locks or unlocks. I derive the (1) number of phone locks and unlocks, and (2) the average duration of time between phone locks and unlocks each day.

Desk Activity The Bluetooth beacons in conjunction with the smartphone application captured the presence of individuals (e.g., at work/home locations). I derive several daily features about activity patterns at work and home each day, including (1) time at work, (2) minutes at desk, (3) mean desk session duration, (4) median desk session duration, (5) percent of time at work spent at desk, (6) and the percent of time of the 24 hour day spent at work. I also compute *break session* information, i.e. the intervals between the participant's desk beacon being out of range and the desk beacon appearing within range. Specifically, I compute daily counts of break sessions at three different interval measures: (7) number of 5-minute breaks, (8) number of 15-minute breaks, and (9) number of 30 minute breaks.

Sleep The wearable sensed the sleep activity of the individuals [234]. Wearables can accurately detect sleep [341, 686]. This measurement was improved by further accounting for phone use and wearable-derived bed times, wake times, and sleep duration drawing on Martinez et al. [400]. In addition to collecting daily measures of (1) bed time, (2) wake time, and (3) sleep duration using this method, I also derive duration measures directly from the wearable for (4) light sleep, (5) deep sleep, and (6) Rapid Eye Movement (REM) sleep [396]).

To compute physical sensor features for clustering the participants into different behavioral contextualizations, I calculate the mean (μ) and standard deviation (σ) of each daily measure described above, for each individual. In addition to μ and σ features, I also compute features characterizing the *regularity* of each measure using Recurrence Quantification Analysis (`RQA`) [660]. `RQA` estimates the number and duration of occurrences of a dynamical system presented through a phase space trajectory [660]. Regularity measures derived from multimodal sensing data have shown valuable utility in predicting psychological traits [653].

In particular, I obtain features using RQA for the *recurrence rate*, which represents the probability that a specific state will occur, and can be interpreted as the repetitiveness of the elements in a given sequence (i.e. the repetitiveness of values across the days for which data was collected). `RQA` is computed using three parameters: (1) the delay parameter τ, which is the delay unit by which the series is lagged, the dimension embedding D, which is the number of embedding dimensions for phase reconstruction, i.e. the lag intervals, and the radius R, which is the threshold cut-off constant used to determine if two points are recurrent or not. For each daily sensor measures series, I use the method recommended by Wallot [649] to compute the optimal parameters for each series, I computing the optimal parameters for each individual, and then using the mean value from this distribution of parameters to apply to the sensor measure stream. Among the `RQA` features, I could not attain useful features for mean and maximum average distance from home, as these `RQA` features show almost no variability across the participants, and hence, I discard them from the final feature set. From mean, standard deviation, and `RQA` aggregation methods, I obtain a total of 76 behavioral features for all participants.

Deriving Features from Social Media Dataset

Longitudinal Social media data of individuals is self-recorded in naturalistic settings. This data also enables us to obtain historical behavior of participants, i.e., from before study participation. Drawing on prior work [162, 166, 531, 533, 564], I obtain a variety of features from the Facebook data of the participants, and summarize them below.

Psycholinguistic Attributes A number of prior work in the space of social media and psychological wellbeing [166, 564] have used psycholinguistic attributes in building predictive models. On the Facebook posts of the individuals, I use the well-validated Linguistic Inquiry and Word Count (LIWC) lexicon [613] to extract a variety of psycholinguistic categories (50 in total). These categories consist of words related to 1) *affect* (categories: anger, anxiety, negative and positive affect, sadness, swear); 2) *cognition* (categories: causation, inhibition, cognitive mechanics, discrepancies, tentativeness); 3) *perception* (categories: feel, hear, insight, see); 4) *interpersonal focus* (categories: first person singular, second person plural, third person plural, indefinite pronoun); 5) *temporal references* (categories: future tense, past tense, present tense); 6) lexical density and awareness (categories: adverbs, verbs, article, exclusive, inclusive, negation, preposition, quantifier); 7) *personal and social concerns* (categories: bio, body, death, health, sexual, achievement, home, money, religion, family, friends, humans, social).

Open Vocabulary n-grams Open-vocabulary based approaches can infer psychological constructs of individuals [564]. I obtain the top 5000 n-gram ($n = 1, 2, 3$) from the dataset as features.

Sentiment An important dimension in the language expressed on social media is the tone or *sentiment* of a social media post, which has also been used to understand psychological constructs and shifts in mood of individuals [252, 529]. I use the Stanford CoreNLP library's deep learning based sentiment analysis tool [393] to identify the major sentiment of a post among positive, negative, and neutral sentiment labels.

Latent Lexico-Semantics (Word Embeddings) Word embeddings are vector representations of language in latent semantic dimensions, enabling us to capture the lexico-semantics of language on social media. Prior work reveals that word embeddings can improve several natural language analysis and classification problems [155, 476, 544]. I use pre-trained word embeddings (GloVe [476]) on an internet corpus of 6B tokens in 50-dimensions to characterize the social media posts of the participants in a 50-dimensional feature space.

Social Capital Social capital is an important aspect and contributor in shaping the lives and behaviors [599]. Drawing on prior work [164], I obtain features quantifying social capital of the individuals based on social media interactions and engagement. I use regular expression based pattern matching to identify individuals' updates relating to 1) check-ins to places (or locations visited), 2) posts of status updates, 3) upload of media (photo or video), 4) spend time (or an occasion) with other people (or friends), 5) change in relationship status and 6) use of apps (such as games or quizzes on Facebook). For each of these social attributes, I compute the number of updates, frequency of updates, and the number of likes and comments received in them.

In total, 5,127 derived features are obtained corresponding to each participant on their social media data.

6.1.3 Aim 1: Contextualizing and Predicting Psychological Constructs

This study focuses on predicting psychological constructs with social media data. Social media use and expressiveness may not only vary significantly across individuals, but also are also driven by offline factors. Therefore, contextualizing on offline behaviors may make models better adapted to the social media signals predictive of psychological constructs per individual. As briefly introduced before, I take a middle-ground approach between fully individualized and fully generalized prediction models, which aims to capture between-individual homogeneity and within-individual heterogeneity. Intuitively speaking, given the sparsity of social media data, for an individual, whose social media data of certain behaviors or moments is "missing" (or lack of within-individual heterogeneity in data), we could fill these gaps by capturing the data from other similar individuals (between-individual homogeneity); here the similarity is captured via offline behavioral clustering.

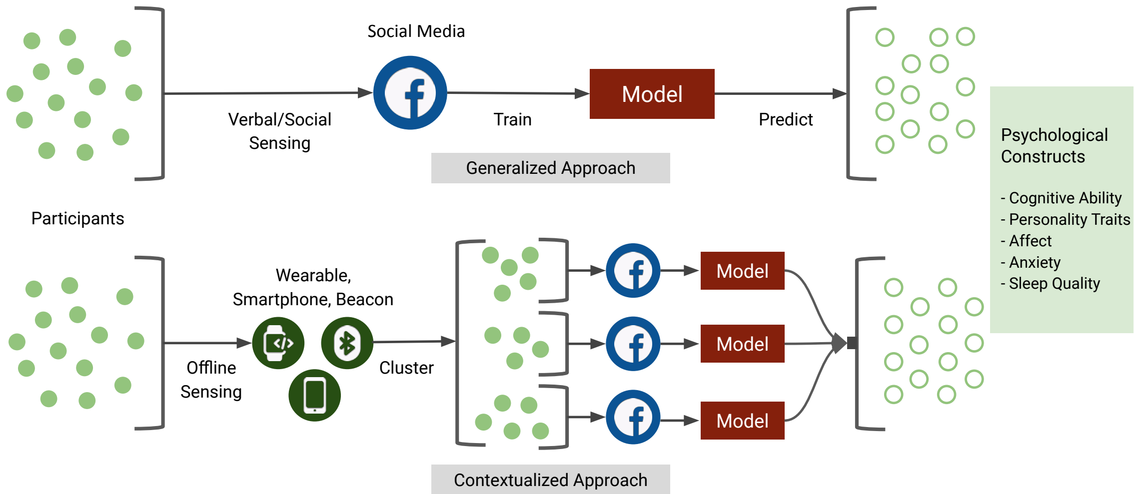


Figure 6.2: A Schematic diagram comparing generalized prediction with social media data and the person-centered approach by contextualizing offline behaviors.

To do so, I cluster individuals on the basis of offline physical behaviors (e.g. sleep, work, phone use, physical activity) captured from sensors on Bluetooth beacons, wearables, and smartphones. Then, I build cluster-specific prediction models of psychological constructs, where each cluster-specific model uses the social media data of participants only within the corresponding cluster. Figure 6.2 schematically summarizes the contextualized prediction approach in comparison to generalized prediction approach with social media data.

Clustering Individuals on Physical Sensor Behavior

To categorize the participants into different clusters based on their behavioral features, I first perform data imputation and feature selection to reduce missing values and feature redundancy. Feature selections and feature transformations are important preliminary steps for any machine learning problem to overcome problems of multi-collinearity, co-variance, etc. among the features — issues that can potentially affect downstream prediction problems [186]. In particular, because the features are derived from multimodal data streams, there is a high likelihood many of the features are already related, are redundant, and/or show extremely high variance and lack predictive power [497]. For example, the activity and stress-related features as captured by the wearable, are both intuitively and theoretically correlated [234].

Starting with 76 physical sensor features for 704 participants in the training and validation data as explained in the previous section, I impute any existing missingness (as is common in any longitudinal and large-scale data collection) in the data by using mean imputation per feature. Next, I conduct step-wise removal of multi-collinear features by calculating the variance inflation factor (VIF) of features against each other [154, 419]. I eliminate correlated features that show a VIF higher than 5, reducing our feature set from 76 down to 46 features.

Building Clustering Models On the above finalized training and validation dataset, I apply four clustering algorithms to obtain the optimal arrangement: K-means, partitioning around medoids (PAM) [632], and two versions of hierarchical clustering. While both hierarchical clustering methods use Wards method for agglomeration between clusters [656], the first method ($hclust_1$) uses Ward's approach on the two observations and/or clusters which were recently merged when updating the distance matrix, while the other method ($hclust_2$) uses Ward's approach on all observations in the merged clusters, i.e. using less shortcuts when updating the distance matrix.

For each clustering algorithm, I test cluster arrangements ranging from 2 to 8 clusters, using the mean Silhouette score [521], Dunn index [190], and the connectivity of the clusters (i.e. the degree of connection within the clusters, measured by k-nearest neighbors) [288] to determine the most optimal clustering arrangement. Table 6.2 presents the results of the clustering tests by varying parameters. I find that the most optimal clustering arrangement in terms of mean Silhouette score and Dunn index is $hclust_2$ with 2 clusters, however, a close second is $hclust_2$ with 3 clusters, where the connectivity score is slightly higher but the mean silhouette score is slightly lower (0.26 vs. 0.27). As the primary research goal is to investigate the utility of building separate social media based models based on *contextualized* behavioral information about individuals, rather than a rigorous evaluation the most optimal

Table 6.2: Comparing goodness of fit metrics for clustering methods and parameters (Number of clusters, Silhouette score, Dunn index, and Connectivity). The **green** highlighted row is the finally used clustering approach.

# Clusters	Sil. Score	Dunn Idx.	Connectivity
K-Means			
2	0.11	0.09	381.27
3	0.04	0.09	431.09
4	0.05	0.09	463.40
5	0.06	0.10	446.15
6	0.04	0.08	576.84
7	0.05	0.08	609.36
8	0.05	0.07	618.97
Partitioning Around Medoids (PAM)			
2	0.03	0.09	223.37
3	0.02	0.09	454.82
4	0.03	0.09	473.11
5	0.02	0.09	521.16
6	0.01	0.09	584.95
7	0.02	0.09	585.91
8	0.01	0.09	599.11

# Clusters	Sil. Score	Dunn Idx.	Connectivity
Hierarchical 1 (distance on recent two observations)			
2	0.23	0.13	214.92
3	0.02	0.10	406.09
4	0.02	0.10	443.02
5	0.03	0.10	453.82
6	-0.03	0.05	584.50
7	-0.03	0.05	626.90
8	-0.02	0.05	632.54
Hierarchical 2 (distance on all observations)			
2	0.27	0.14	149.10
3	0.26	0.14	151.66
4	0.05	0.10	350.83
5	0.03	0.10	454.58
6	0.04	0.10	455.95
7	0.04	0.10	459.50
8	0.04	0.10	462.36

clustering arrangement of the dataset, I consider that having more individualization in behavioral categories might provide a better evaluation of the theoretical approach. I proceed with the analysis and model building using the $hclust_2$ clustering algorithm with 3 distinct clusters.

Describing Clusters on Physical Behaviors Applying the *hclust2* clustering algorithm with three distinct clusters to the training and validation subset, we find Cluster $\mathtt{C}_1$ to have the majority of the participants (N=601, 85%), while Cluster $\mathtt{C}_2$ has N=76 (11%) participants, and Cluster $\mathtt{C}_3$ has N=27 (4%) participants. To better understand how the clusters differed among the behavioral features we generated, we apply the Kruskal-Wallis H-test to each of the 46 behavioral features with the responding variable as the behavioral feature, and the independent variable as the cluster membership category. We use the Kruskal-Wallis H-test as the cluster sizes are very different in size, and therefore we cannot likely assume a normal distribution of the feature values within each cluster. Table 6.3 reports the top 20 behavioral features with significant H-statistic values from the tests.

We find many regularity (RQA-based) features in the top 20 features. Regularity in minutes at desk per day, desk session duration, REM sleep duration, and number of phone

Table 6.3: Mean z-scores per cluster for top 20 significant features as per Kruskal Wallis H-test used for clustering. Statistical significance reported after Bonferroni correction (*** $p < .001$, ** $.001 < p < .01$, * $.01 < p < .05$).

Features	C_1	C_2	C_3	H-stat.
Phone Use				
Regularity of Number of Phone Unlocks per day	-0.15	1.09	0.33	63.12***
Regularity of Duration Spent with Phone Unlocked per day	-0.15	1.12	0.22	38.38***
Mean Number of Phone Unlocks per day	0.06	-0.38	-0.19	20.83**
Work Behaviors				
Regularity of Minutes at Desk per day	-0.17	0.12	3.46	79.72***
Regularity of Mean Desk Session Duration	-0.14	0.10	2.88	74.34***
Regularity of Median Desk Session Duration	-0.14	0.02	2.95	60.73***
Regularity of Percent Time Spent at Work per day	-0.14	0.03	3.07	70.84***
Mean Percent of Time at Work Spent at Desk	0.08	-0.03	-1.55	47.62***
Stdev. of Time at Work Spent at Desk	0.07	-0.11	-1.24	20.00**
Sleep				
Regularity of Total REM Sleep Duration per night	-0.18	1.44	-0.10	76.27***
Mean of Total REM Sleep Duration per night	0.13	-0.91	-0.32	54.48***
Stdev. of Total REM Sleep Duration per night	0.14	-0.98	-0.24	37.68***
Regularity of Total Deep Sleep Duration per night	-0.17	1.26	0.13	32.52***
Regularity of Nightly Bed Time	-0.03	0.38	-0.33	32.03***
Mean of Total Light Sleep Duration per night	0.11	-0.89	0.02	27.60***
Stdev. of Nightly Bed time	-0.07	0.23	0.88	18.07**
Physical Activities				
Regularity of Steps Count per day	-0.11	0.85	-0.01	51.58***
Regularity of Total High/Strenuous Activity Duration per day	-0.14	1.08	-0.06	46.92***
Regularity of Total Activity Duration per day	-0.14	1.13	-0.03	30.70***
Mobility				
Mean of Total Distance Travelled per Day	0.05	-0.33	-0.41	14.55*

unlocks are strong explanatory features to distinguish the three clusters. To investigate more specifically how the top features differ across each cluster, we transform the values for these features into z-scores within the entire participant set — Table 6.3 also provides the mean z-scores. z-score transformations are not sensitive to absolute values and measure the raw value in terms of standard deviations above or below the mean. I observe differences in C_3 compared to C_1 and C_2 for a number of the features, primarily with respect to work behaviors. Participants in C_3 had much higher daily regularity in minutes at their work desk per day (mean z=3.46) than those in C_1 (mean z=−0.17) or C_2 (mean z=0.12), but also on average spent a much lower percentage of their workday at their desk (mean z=−1.55) compared to those in C_1 (mean z=0.08) or C_2 (mean z=−0.03). I also observe distinct differences in C_2 compared to C_1 and C_3 with respect to sleep patterns. Participants in C_2 had more regularity in nightly seconds of REM sleep (mean z=1.44) compared to C_1 (mean z=−0.18) or C_3 (mean z=−0.10), but also had a lower average in nightly seconds of REM sleep (mean z=−0.91) compared to C_1 (mean z=0.13) or C_3 (mean z=−0.32).

Examining Clusters On Demographic Composition I examine the demographic composition of each cluster. Creating separate participant clusters based on behavioral features might be similar to directly clustering on demographic information. For instance, older adults are known to be more sedentary [88], and factors like age and gender have been shown to explain daily smartphone usage [21]. As this study focuses on the utility of building person-centred models based on passively sensed behavioral data, rather than static demographic information, it strives to have the clusters to have heterogeneous demographic compositions.

To test the heterogeneity of the demographic composition across clusters, I perform χ^2 tests between the clusters and the categorical demographic variables (Gender, Education Level, Income, Tenure, and Supervisory role), and a one-way ANOVA test between the clusters and age (the only numeric demographic variable). The tests reveal no significant association between the clusters and Age (χ^2(2)=0.91, p=0.63), Gender ($\chi^2_{Pearson}$(2)=3.06, p=0.22), Income ($\chi^2_{Pearson}$(12)=10.99, p=0.53), Supervisory Role ($\chi^2_{Pearson}$(2)=3.72, p=0.16), and Tenure ($\chi^2_{Pearson}$(18)=19.97, p=0.34). While this shows a weak significant association between the clusters and Education ($\chi^2_{Pearson}$(8)=17.25, p=0.03), the effect size is very small ($\hat{V}_{Cramer}$=0.08). I observe slight compositional differences in Education in C_3, such that there are proportionately more participants with High school as the highest level of education (7%), compared to C_1 (1%) and C_2 (0%). C_3 also has proportionately less participants with a College degree as the highest level of education (44%), compared to C_1 (55%) and C_2 (54%). However, these significant demographic differences are relatively negligible and only occur for education. Therefore, I conclude that the clusters are much more strongly separated by the passive behavioral features than by demographic information.

Predicting Psychological Constructs with Social Media

I use the features described in Section subsection 6.1.2 to predict self-reported psychological constructs (Table 6.1). For each psychological construct, I build two kinds of models: 1)

generalized models which are built on the entire dataset of all participants, 2) **contextualized models** which are separately built per behaviorally contextualized clusters. Here, the generalized prediction models emulate typical practices of predicting behavioral attributes with social media data, whereas the clustered models are more person-centered driven by incorporating people's offline behaviors (passively inferred via physical sensors).

I use k-fold cross-validation (k=5) for parameter tuning and evaluation with *pooled accuracy technique* on the training and validation subset of the data, i.e., for each model, I first divide the dataset into five equal segments, then iteratively train models on four of the segments to predict on the held-out fifth segment, and finally collate all the predicted values together and compare their collated against actual values using Pearson's correlation (r) and Symmetric Mean Absolute Percentage Error (SMAPE). I adopt several prediction algorithms spanning across linear regression (with and without L_1, L_2 regularization), gradient boosted random forest (GBR), support vector regressor (SVR), and multilayer perceptron (MLP).

I also transform the social media feature set using Principal Component Analysis (PCA) with a singular value decomposition solver, selecting the number of components on the basis of explained variance [297][1]. However, prediction models using PCA transformed features show no improvement over those using raw features (no-PCA transformation), likely because language and n-gram features are inherently sparse and contain predictive information despite the variance and sparsity. The remaining paper concerns analyses with raw features, which serves an additional advantage of feature interpretation and model explanation with respect to contextualization.

To verify that the training models do not overfit, I also apply the cross-validated and trained models to the held-out unseen subset of the data (N=50) to test performance on unseen data (introduced at the beginning of Section subsection 6.1.2). I derive the 46 physical sensor features for the held-out data, and apply the same trained *hclust*$_1$ model to obtain cluster

[1]The Appendix Table B.3 and Table B.4 show the performance of modeling with PCA-transformed features. Qualitatively, these predictions show similar comparison directions between generalized and contextualized models as observed in models with non-transformed features.

labels for the held-out data. I again evaluate the relative performance of the generalized and contextualized models on the held-out data.

6.1.4 Aim 2: Evaluating Performance of Contextualized and Generalized Models

On the best performing algorithm for generalized and contextualized prediction models from the above, I compare the performance metrics of these predictions for the cross-validated evaluation with the training data (Table 6.4, detailed metrics in Appendix, Table B.1 and Table B.2), and for the held-out data (Table 6.5). To measure statistical significance in prediction differences, I conduct t-tests using the dependent overlapping correlation method, which controls for comparing against a common variable of interest (here, each psychological construct) [191]. I observe that the efficacy of contextualizing on offline behavior for social media based predictions can be explained by the theoretical associations between the construct and its offline manifestations. I now discuss the performance of the models, and ask: When does contextualization help? Are there any cases where contextualization does not improve over generalized models?

Cognitive Ability

I use the Shipley scales to obtain ground-truth measures of two kinds of cognitive ability. The Shipley (Abstraction) scale measures fluid cognitive ability, which is how one thinks logically, reasons, and problem solves in novel situations. The Shipley (Vocabulary) scale measures crystallized cognitive ability [578], or an individual's grasp of general and cultural knowledge including verbal communication [103]. These two abilities mutually interact and combine to form overall individual cognitive ability [313].

Table 6.4 and Table 6.5 compare the best predictions as per generalized and clustered models in the cross-validated evaluation and held-out data respectively. In the case of abstraction, there is no significant difference in the performances of generalized and clustered models. However, in the case of vocabulary, there is a statistically significant difference

Table 6.4: Cross-validated Evaluation: Comparing the accuracy metrics of best performing generalized and contextualized prediction models. Statistical significance is computed using t-test as per dependent overlapping correlations [191] on predictions by generalized and contextualized models for each construct. For significant rows, **pink** bars indicate a **decrease in performance** in contextualized models compared to generalized models and **green** bars indicate an **increase in performance** (*** p<.001, ** .001<p<.01, * .01<p<.05).

Construct	Generalized		Contextualized		Comparison		
	r	SMAPE	r	SMAPE	Δr %	ΔSMAPE %	t-stat.
Cognitive Ability							
Shipley (Abstraction)	0.25	6.81	0.23	6.88	-8.00	1.03	-1.73^{-}
Shipley (Vocabulary)	0.29	4.13	0.21	4.25	-27.59	2.91	-4.82***
Personality Traits							
Openness	0.25	6.89	0.29	6.08	14.81	-11.76	1.94*
Conscientiousness	0.13	7.29	0.19	7.08	46.15	-11.76	2.80**
Extraversion	0.17	8.54	0.21	8.46	23.53	-0.94	1.70*
Agreeableness	0.17	5.84	0.19	5.89	11.76	0.86	0.88^{-}
Neuroticism	0.12	13.56	0.18	13.09	50.00	-3.47	2.50*
Affect and Wellbeing							
Pos. Affect	0.13	7.10	0.14	6.90	7.69	-2.82	0.56^{-}
Neg. Affect	0.11	10.90	0.13	10.89	18.18	-0.09	-1.13^{-}
Anxiety (STAI)	0.12	9.66	0.21	8.51	75.00	-11.90	5.61***
Sleep Quality (PSQI)	0.15	16.02	0.25	10.59	66.67	-33.90	5.07***

(t=-4.61), where the generalized model performs 27.6% better in r and 2.41% better in SMAPE in the cross-validated evaluation. I observe similar prediction results in the held-out data, where the generalized model performs 29.41% better in r and 8.41% better in SMAPE. However, the difference in the held-out data is not quite significant (t=-0.99, p=0.08), likely due to the smaller sample size.

The above suggests that clustering individuals on physical and offline behaviors does not add any new information in predicting cognitive ability. I construe that more heterogeneity of individuals in training sample and larger size of data are in fact stronger factors in predicting cognitive ability, likely because language is known to be a correlate of cognitive ability, more strongly in the case of crystallized cognitive ability (vocabulary) [556].

Personality Traits

Personality traits are considered to be robust and parsimonious correlates of a variety of individual outcomes, characteristics, behavior, and abilities [355]. I find that in both the cross-validated evaluation and the held-out data, contextualized predictions perform significantly better for Openness, Extraversion, and Neuroticism. Contextualized predictions

Table 6.5: Held-out data: Comparing the accuracy metrics of best performing generalized and contextualized prediction models. Statistical significance is computed using t-test as per dependent overlapping correlations [191] on predictions by generalized and contextualized models for each construct. For significant rows, **pink** bars indicate a **decrease in performance** in contextualized models compared to generalized models and **green** bars indicate an **increase in performance** (*** p<.001, ** .001<p<.01, * .01<p<.05).

Construct	**Generalized**		**Contextualized**		**Comparison**		
	r	SMAPE	r	SMAPE	Δr %	ΔSMAPE %	t-stat.
Cognitive Ability							
Shipley (Abstraction)	0.36	4.52	0.33	5.03	-8.33	11.28	-0.36^{-}
Shipley (Vocabulary)	0.34	4.65	0.24	5.04	-29.41	8.40	-0.99^{-}
Personality Traits							
Openness	0.51	5.00	0.79	4.23	66.67	-15.40	4.00***
Conscientiousness	0.19	6.87	0.21	6.06	10.53	-11.79	1.26^{-}
Extraversion	0.19	7.71	0.32	6.88	68.42	-10.77	2.57**
Agreeableness	0.38	6.51	0.62	5.81	63.16	-10.75	3.77***
Neuroticism	0.12	12.86	0.63	11.11	425	-13.61	7.95***
Affect and Wellbeing							
Pos. Affect	0.30	9.20	0.60	8.54	100	-7.17	2.57***
Neg. Affect	0.20	11.33	0.42	10.87	110	-4.06	2.25***
Anxiety (STAI)	0.14	11.57	0.33	8.78	135.71	-24.11	1.14*
Sleep Quality (PSQI)	0.16	16.49	0.41	12.33	156.25	-25.23	2.27*

of Conscientiousness perform significantly better in cross-validated evaluation, (t=2.80) without a significant difference in performance in the held-out set (t=1.26, p=0.21). Conversely, contextualized prediction for Agreeableness does not perform significantly better in the cross-validated evaluation (t=0.88, p=0.38), but prediction is significantly better in the held-out set (t=3.70). Despite these inconsistencies in statistical significance, there is a general trend towards increased performance in contextualized predictions for Conscientiousness and Agreeableness. The inconsistencies in statistical significance for improved Agreeableness predictions with contextualized models partially aligns with prior meta-analysis of physical activities and personality traits that found Agreeableness to show the least significant relationship with physical activities [505]. Note that the participant pool is drawn from information workers, and certain activities such as work behaviors and phone use have been found to be direct correlates of traits like Neuroticism [154], so capturing such information during clustering (Table 6.3) may have contributed to the effectiveness of person-centered models. I construe that driving contextualized models through physical behavior based clusters likely allows to capture distinct linguistic features per cluster that are better predictive of personality traits.

Affect and Wellbeing

The contextualized models reveal no benefit for predicting positive or negative affect in the cross validation data set, though a significant benefit is found in the held out data.. This inconsistency suggests that there is inconclusive evidence whether clustering individuals on their physical activities or offline behaviors contributes new information in the prediction of affect. It is likely that offline behaviors might not provide enough new information for predicting more moderate constructs of day-to-day affect like those measured with the PANAS scale.

However, for anxiety and sleep quality (PSQI), there are large and significant improvements for the contextualized predictions over generalized models in both cross-validated and held-out evaluations. I conjecture that these improvements are due to strong correlations between physical behaviors and sleep quality and anxiety. For instance, the duration of deep sleep is known to have a significant effect on reported sleep quality [116], while improved physical activity has a long-established relationship in reducing subjective anxiety [60]. Although extreme affective disorders like depression (which is often comorbid with anxiety) [253] are known to have relationships with offline behaviors like sleep and mobile phone use [616], offline behaviors might not provide enough new information for predicting more moderate constructs of day-to-day affect like those measured with the PANAS scale. I note that the clustering approach includes features obtained using wearables and smartphones, both of which capture behaviors correlated with sleep (e.g., accelerometer data, phone use, etc.), likely helping the contextualized predictions of sleep quality.

Robustness of Contextualized Person-Centered Approach

I test for empirical robustness of the contextualized person-centered approach against typical approaches of using physical sensor features for prediction (M_s models), as well as using all (physical sensor and social media) features together, i.e., a multisensor feature fused model (M_{ms} models). I conduct similar rigorous model and parameter tunings as above, and obtain

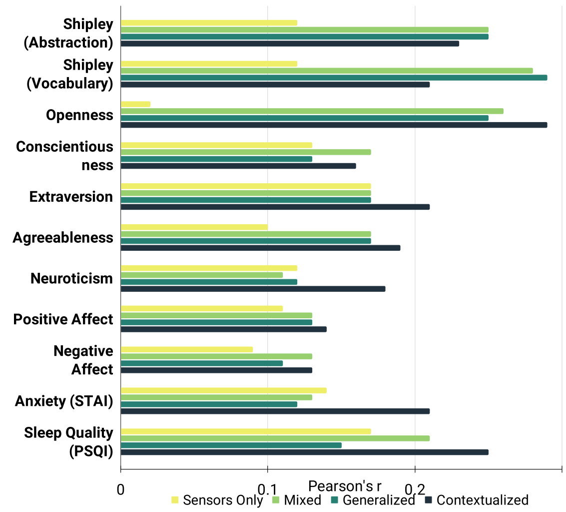


Figure 6.3: Comparing performance (r) various modeling approaches for predicting psychological constructs.

the best models of predicting the constructs. Appendix Table B.5 and Table B.6 present the detailed prediction results for M_s and M_{ms} models respectively, and Figure 6.3 presents a summary view of prediction performance comparison across various modeling approaches. For the M_s models, the prediction performance is considerably worse than contextualized models for all psychological constructs, with the strongest prediction from any M_s model is for Extraversion (r=0.17, SMAPE=8.41).

I find that the best M_{ms} models of cognitive ability perform as similar as generalized social media models, and better than contextualized models for both abstraction (r=0.25, SMAPE=6.66) and vocabulary (r=0.28, SMAPE=4.10). For personality traits, I find that M_{ms} performs similarly or worse than contextualized models in openness (r=0.26, SMAPE=6.40), conscientiousness (r=0.17, SMAPE=7.86), extraversion (r=0.17, SMAPE=8.35), agreeableness (r=0.17, SMAPE=5.91), and neuroticism (r=0.11, SMAPE=12.93). For affect and well-being, M_{ms} performs similarly in positive (r=0.13, SMAPE=6.77) and negative (r=0.13, SMAPE=11.18) affect, and significantly worse in anxiety (r=0.13, SMAPE=16.07) and sleep quality (r=0.21, SMAPE=14.02). Together, this suggests that the approach of person-centered contextualization not only mines signals in the multimodal sensing data better, but also likely filters out noisy information from user groups whose behavior is too far away from the average group. Further, the person-centered contextualization approach provides additional theoretical interpretation and explanation which I elaborate further in the following sections.

In addition, I also target rejecting the null hypothesis that any prediction improvement by the contextualization approach is by chance or any random cluster-label assignment. Drawing on permutation test approaches [17, 537], I permute (randomize) the cluster label of all individuals, and repeat the entire pipeline predicting of psychological constructs. I run 1,000 such permutations, and I find that the probability (p-value) of improvement by a random-cluster assignment over contextualized approaches is almost zero across all the measures (p=0.002 for abstraction, p=0.001 for positive affect are the only non-zero probabilities). This rejects the null hypothesis and provides additional statistical significance and credibility to the person-centered approach of contextualization using offline behavioral clustering.

6.1.5 Aim 3: Offline Contextualization and Social Media Language

The sensitivity of social media data to people's unique characteristics and variable social media use motivates us to study person-centered contextualized predictions. I have already proposed and validated an approach to contextualize social media predictions of psychological constructs by clustering people on offline behaviors (Aim 1 and 2). Next, I interpret the same clusters in terms of people's social media use. Our third research aim targets understanding how the social media language varies by the clusters. I investigate if clustering individuals on *offline* behaviors leads to clusters of individuals who also have different online behaviors.

To understand the language differences better, I first interpret the composition of clusters on psychological constructs which can help us validate the theoretical foundation of building person-centered models on contextualized offline behaviors. Then, for each cluster, I obtain salient language use and interpret that with respect to cluster composition in offline behaviors,

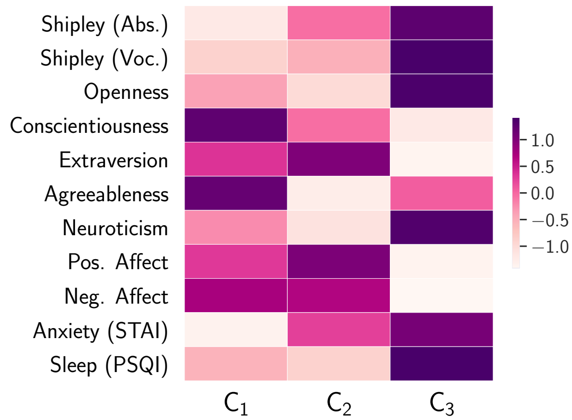


Figure 6.4: Heatmap representing the clusters on mean psychological constructs. Values are z-transformed per measure.

psychological constructs, and the literature.

Interpreting Cluster Composition on Psychological Constructs

I examine the between-cluster differences in psychological constructs (or the outcome measures). Figure 6.4 shows z-transformed representation of the mean composition of each cluster per construct. At a mean-aggregated level, $\mathtt{C}_1$ shows the greatest average conscientiousness (μ=3.89, σ=0.66) and agreeableness (μ=3.90, σ=0.57). $\mathtt{C}_2$ shows greatest average positive affect (μ=35, σ=5.89), negative affect (μ=17.49, σ=4.97), and anxiety (μ=38.76, σ=10.49). $\mathtt{C}_3$ shows greatest average cognitive ability in both abstraction (μ=17.41, σ=2.61) and vocabulary (μ=33.41, σ=3.84), openness (μ=3.9, σ=0.41), neuroticism (μ=2.49, σ=0.76), and self-reported sleep (μ=7.03, σ=2.92), while showing low affective traits.

For all measures, Kruskal-Wallis H-tests across clusters show no statistical significance, which could mean that each cluster is already composed of heterogeneous psychological traits. This within-cluster variation in psychological constructs suggests that clustering on offline (dynamic) behaviors does not necessarily translate to clustering individuals with only "similar" psychological constructs. For each cluster and psychological construct, I compute

Table 6.6: Comparing psycholinguistic attributes across clusters. Statistical significance reported after Bonferroni correction (*** $p < .001$, ** $.001 < p < .01$, * $.01 < p < .05$).

Category	$\mathbf{C_1}$	$\mathbf{C_2}$	$\mathbf{C_3}$	**H-stat.**
Affect				
Anger	0.04	0.41	0.29	46.48***
N. Affect	0.15	0.26	0.39	28.46***
P. Affect	0.32	0.62	0.55	105.52***
Sadness	0.14	0.21	0.29	18.42***
Cognition				
Causation	0.19	0.42	0.25	53.01***
Certainty	0.11	0.54	0.19	81.71***
Cog. Mech.	0.51	0.69	0.55	46.03***
Inhibition	0.27	0.41	0.20	22.11***
Discrepancies	0.15	0.48	0.27	81.84***
Tentativeness	0.24	0.59	0.30	65.10***
Perception				
Feel	0.14	0.40	0.46	79.17***
Hear	0.11	0.41	0.31	64.18***
Insight	0.22	0.57	0.41	74.12***
Percept	0.20	0.55	0.57	94.32***
See	0.10	0.34	0.59	90.77***

Category	$\mathbf{C_1}$	$\mathbf{C_2}$	$\mathbf{C_3}$	**H-stat.**
Interpersonal Focus				
1st P. Sing.	0.20	0.26	0.43	23.74***
1st P. Plu.	0.05	0.37	0.42	78.10***
2nd P.	0.12	0.39	0.27	77.64***
Indef. P.	0.35	0.40	0.54	19.92***
Lexical Density and Awareness				
Adverbs	0.30	0.52	0.54	86.57***
Article	0.28	0.64	0.49	110.01***
Verbs	0.42	0.59	0.61	64.05***
Aux. Verbs	0.25	0.57	0.54	98.30***
Conjunction	0.27	0.63	0.61	107.54***
Exclusive	0.33	0.52	0.45	39.65***
Inclusive	0.30	0.67	0.58	104.74***
Preposition	0.37	0.58	0.77	111.5***
Negation	0.09	0.21	0.19	51.07***
Quantifier	0.16	0.52	0.17	66.07***
Relative	0.22	0.71	0.56	127.55***

Category	$\mathbf{C_1}$	$\mathbf{C_2}$	$\mathbf{C_3}$	**H-stat.**
Temporal References				
Future Tense	0.10	0.31	0.58	100.78***
Past Tense	0.18	0.34	0.37	73.22***
Present Tense	0.34	0.54	0.66	79.73***
Personal and Social Concerns				
Achievement	0.24	0.52	0.45	82.62***
Bio	0.08	0.31	0.54	98.16***
Body	0.02	0.29	0.38	78.54***
Family	0.09	0.13	0.22	28.11***
Friends	0.05	0.19	0.24	69.43***
Health	0.14	0.30	0.28	33.43***
Home	0.05	0.14	0.14	69.07***
Humans	0.08	0.19	0.49	81.18***
Money	0.06	0.46	0.30	83.66***
Religion	0.08	0.20	0.19	14.02***
Social	0.33	0.60	0.46	81.78***
Work	0.09	0.31	0.32	107.08***

the coefficient of variation (cv) [208], expressed as a percentage in the ratio of standard deviation to mean of a distribution, and quantifies the amount of variability with respect to the mean of the distribution — higher values indicate higher variability. The cluster-wise cv is on an average 7% lower compared to the entire (non-clustered) data's cv across all the measures. Together, this suggests that clustering finds a compromise in both preserving and reducing the variability in training data compared to the entire data. Methodologically, the within-cluster heterogeneity in outcomes plausibly helps the data within-clusters to be neither too biased (if only predicting homogeneous or skewed distribution of labels), nor too varianced (predicting high variance distribution of label, e.g., the entire data), thus helping the predictive performance of psychological attributes, as noted in Section subsection 6.1.4.

Examining Social Media Language Differences across Clusters

Next, I investigate how offline contextualization leads to groups of individuals with different social media use, and what are the likely theoretical interpretation of such groupings in terms of understanding psychological constructs. I base this examination on the differential psycholinguistic usage of individuals. I measure the statistical difference in language using Kruskal-Wallis H-test. Table 6.6 characterizes the clusters on psycholinguistic use categories which show significant differences as per Kruskal-Wallis H test. The average standard deviations across the categories are 0.12 for C_1, 0.16 for C_2, and 0.16 for C_3. I discuss this below.

Although **Cluster $\mathbf{C_1}$** is the largest (468 individuals), it shows the lowest standard deviation (0.12), suggesting that this cluster is psycholinguistically least heterogeneous. C_1 shows the lowest use of all psycholinguistic attributes, suggesting that these individuals are typically less expressive on social media. An alternate explanation would be C_1 represents the more common language use on social media — language containing lesser presence of non-content words (articles, prepositions, etc.) that is known to be associated with less complex language [613]. Moreover, the low mean in psycholinguistic attributes could also be associated with high mean conscientiousness of C_1 (Figure 6.4), aligning with prior findings on social media language and psychological traits [564]. Similarly,examining C_1's offline behaviors (Table 6.3) shows these individuals typically score low on the regularity based features (work behaviors, sleep, and physical activities), suggesting that the relative heterogeneity of offline behaviors, does not translate to online behavior and social media expressiveness. Physically, C_1 travels the most (high mean total distance travelled per day), which suggests interesting associations between their offline mobility and online posting behavior, as has also been noted in prior work [333].

Cluster $\mathbf{C_2}$ shows the greatest use of *anger* and *positive affect*, and all cognitive attributes, suggesting these individuals have an average high emotion on social media language. Interestingly, the same cluster's composition showed high average affect and anxiety traits (Figure 6.4), suggesting that individuals with higher affect traits are likely to be more expressive with affect and emotional language on social media. These individuals have a high use of function words such as *articles*, *auxiliary verbs*, *conjunction, inclusive, exclusive, quantifier, and relative*. Function words are strong linguistic markers of understanding psychological processes [475]. These individuals also show a high use of *achievement* and *money* related language, which may be associated expressiveness about their career and self-actualization.

The observations reveal plausible connections between these psycholinguistic trends and the behavioral sensing features which best separated the clusters (see Table 6.3), e.g., the greatest use of *health* words may be associated with high physical activities shown by them [14]. Additionally, $\mathtt{C}_2$ has an average low duration of REM sleep per night, more regular daily exercise patterns, and more regularity in time spent with their phone unlocked per day, possibly because participants might be more driven for achievement in social, work, or athletic goals, sometimes at the expense of their sleep quality [553].

Cluster $\mathbf{C_3}$ shows the greatest use of pronouns, with pronoun use associated with narrative language and interpersonal discourse on social media [475]. For instance, the greater use of first person singular pronouns (e.g. "I", "me") suggests narrating personal experiences and self-reflection, and that of first person plurals (e.g. "We", "Us") indicates narrating experiences as collective identities [126]. These individuals also score high on the use of language related to social concerns and relationships such as *family, friends, home, and humans*. I again see plausible connections between these psycholinguistic trends and the behavioral sensing features which best separate the clusters. $\mathtt{C}_3$ has on average more regularity in the percent of time they spent at work and desk each day. Regularity in work and at desk suggest these participants might prefer a more regular daily work schedule to balance a need for a more consistent family or social life outside of work.

The above cluster-wise decomposition on social media language reveals how people's offline behaviors can help us group individuals who are also separated in social media language. Further, the analyses in Aim 1 and 2 also reveals how this approach helps improving predictions of psychological constructs, particularly those that bear strong associations with physical behavior (e.g., sleep quality). While it is intuitive that offline dynamic behaviors indeed drive online behavior, there is a paucity of theoretical evidence [14]. This study sheds light on this important aspect and opens up opportunities for future explorations of understanding human behavior.

6.1.6 Discussion

This study adopted machine learning and statistical modeling approaches to contextualize social media predictions of psychological constructs. I first clustered individuals on physical activities as captured by passive sensors and then built cluster-specific prediction models of psychological constructs. The effectiveness of such an approach varies by construct, suggesting that personalization is only better than generalization in specific circumstances. In fact, clusters based on dynamic offline behaviors were not only heterogeneous in static traits (Section paragraph 6.1.3 and subsubsection 6.1.5), but were also separated on social media language use (Aim 3). This study is grounded on the Social Ecological Model that individual behaviors are influenced by factors related to an individual and their context [102]. Beyond just an evaluation of predictive performance, this study provides insights applicable to studies grounded in similar theoretical settings where there is a need to focus on comprehensive social ecological signals, and an opportunity to infer behavioral and psychological attributes of individuals.

Theoretical and Methodological Implications

Beyond Traditional Forms of Personalization Recent research in applied computing has highlighted the value of personalizations via "one-size may not fit all" arguments, as all individuals are not the same and have different experiences [524, 670]. This has also motivated various ubiquitous computing research in personalizing interactive, informatic, and intervention systems [125, 156, 374, 376]. While person-centered analyses have been studied in other disciplines such as social science, psychology, and health [309, 364, 670], such analyses remain under-explored in computational assessments despite the abundance of data. A close application is personalized content recommendation systems [576]. These personalizations have typically relied on a single modality of data (e.g., historical content browsing), along with demographic and static information. Relying on isolated modalities are limited by several blindspots that challenge the comprehensibility of the models, such as the varying data quality across individuals. Importantly, collecting demographic information not only removes user anonymity and threatens data privacy, which is a growing public concern in social media use, but also can promote bias, exclusion, and stereotypes [314]. The implications of this study situate that with recent research revealing that behavioral predictions can sway away from demographic- and static data, by only accounting for short-term and behavioral data [154, 534].

Further, based on the examination of the association between demographic information and target constructs (see Figure 6.1), it is not readily evident that demographic information would provide more accurate predictions. In contrast, this study leverages naturalistic behavior collected via multimodal sensing to guide person-centered analyses. In comparison to stratifications on demographics and static traits, or other forms of strata assumptions, passive sensing allows us to cluster individuals on physical behaviors, which is robust and dynamic. The efficacy of person-centered models is plausibly explained by the notion that sensing streams both independently, as well as in conjunction can predict the constructs in consideration [268, 511, 595, 651, 653]. This study applies new ways of thinking about person-centered approaches in human-centric, context-aware, and social sensing and applications requiring personalized attributes.

Complementary Prediction Approaches this study contributes to the body of literature studying the complementary advantages of variable-centered and person-centered approaches in various social science and psychological constructs [364]. Person-centered approaches allow investigating individual attributes with precision and personalized context-adaptation. The improved predictions are likely due to personalized training datasets by stratifying individuals on their lifestyle and offline behaviors, rather than relatively less helpful demographic information. On the other hand, no difference in performance could be either due to better statistical power of larger training data or due to no added signal in personalized training data. The findings support Howard and Hoffman's study that determined no single approach whether generalized or person-centered can be considered to be the "best", and it depends on the particular problem of interest and research setting (here, contextualizing predictions on offline behaviors) [309].

Trade-off between Statistical Power and Personalization social media data is sensitive to people's self-presentation, context, and other factors, thereby making it harder for generalized prediction models that target behaviors of average populations. Moreover, given that social media data is characteristically sparse (both within and between individuals), it may not be ideal for fully-individualized models. This study overcomes these challenges by using data from offline behaviorally similar individuals, thereby increasing the training data compared to complete personalization while preserving the personalization aspect. However, the training data size in contextualized models is *still smaller* than that of generalized models. Therefore, this study posits that personalized research requires a consideration of the trade-off between statistical power against the personalization component of predictions. This trade-off would likely arise in any kind of personalized predictions, and potentially builds on the classical "bias-variance trade-off" in machine learning predictions [54] — over-personalized models can be too biased, whereas over-generalized models can suffer from variance in the dataset.

Generating New Hypotheses This study allows generating hypotheses on the relationship between human behavior, psychological constructs, and personalized predictions. These hypotheses can guide us to explore newer questions on what factors make some attributes personalizable, and how between-individual homogeneity and within-individual heterogeneity of information can serve as either a noise or a signal in such predictions. Example hypotheses guided by the findings are, 1) social media sufficiently predicts attributes related to cognitive ability, 2) physical behaviors may not be as effective as social media in predicting affect, 3) social media data needs to be complemented with offline data to accurately predict a physical measure such as sleep quality, and so on.

Implications for Researchers and Practitioners

This study demonstrates the efficacy of person-centered predictions of psychological constructs (cognitive ability, personality, affect, and wellbeing) by using complementary ubiquitous technologies. This study takes a critical stance to reflect upon conducting personalized predictions in practice.

Trade-offs between Generalized and Personalized Models Contextualizing social media predictions using passively sensed offline behavior allows us to go beyond the more common, user-profiling like approaches on demographic information based on one's age, race, and gender, which are not only less-robust, but also could lead to misleading findings or "stereotypy" about particular demographic groups. On the other hand, building personalized prediction models with dynamic and mutable behaviors demands additional overhead, including and not limited to obtaining an individual's multiple modalities of data which is both longitudinal and dense.

This study provides insights on what kind of measures may or may not be personalizable. Personalized prediction is considered a useful approach for improving user experience and understanding human behavior in a variety of problem settings, however, personalization is costly in terms of statistical power and effort. One way to navigate through that is to conduct pilot studies with a small number of participants' data and groundtruth measures, before investing and implementing these approaches at scale. Such approaches can identify an appropriate granularity of an effective personalization, and sensors are most likely to improve predictions of a particular construct. Indeed, more data *does not* necessarily lead to better predictions as found in this study. For instance, researchers interested in social media and cognitive ability could use this study as evidence that social media data is sufficient, and predictions would not be improved by additional effort spent collecting passive sensor data and clustering on it.

Considerations for When to Personalize When considering person-centered predictions, one should consider the need and means. Here, social media data is sensitive to an individual's choice of social media use and expressiveness. This study was theoretically motivated in that social media activity is a function of people's offline behaviors and therefore clustered individuals on these behaviors for the analyses. A similar analog of data sensitivity and variability in quality in case of other sensors could be compliance (e.g., use and non-use of a wearable), and attributes that may drive compliance may be accounted for to cluster individuals.

In parallel, the theoretical motivation of personalizing with respect to a construct is also important to consider. For instance, in precision medicine or when dealing with health constructs, it could be critical to conduct personalized assessments, as many health conditions have clinical heterogeneity, driven by varying experiences and traits of people. In such cases, generalized models capturing average target behavior may not provide actionable insights or useful information to build Just In Time Adaptive Interventions (JITAIs) to address such health conditions [376, 593]. Similarly, in the workplace context, when predicting workplace outcomes, it might be useful to incorporate apriori information about the type and hours of work for contextualization.

The efficacy of personalized predictions is also plausibly dependent on the universality and applicability of a psychological construct on given population. For instance, given that sleep is a universal activity across individuals irrespective of their mutable and immutable characteristics, personalizing sleep quality predictions using physical activity turned out to be significantly effective. However, in cases of less-universal or cohort-specific metrics, such as academic success in grade school, it may make more sense to use demographic or other apriori groupings, e.g. grade level. In addition, person-centered approaches may be uniquely valuable in the cases of rare and less-prevalent attributes, such as understanding phobia, anxiety or psychotic disorders, where globalized datasets may be imbalanced and negatively-skewed due to rarity of the condition, significantly impacting the predictions.

This study has an implication regarding *the reasonable use of sensors in accordance with the construct of interest*. As an example, contextualization improved predictions of sleep quality but impaired predictions of cognitive ability. One likely reason that person-centered models predicted sleep quality significantly better was that the multimodal sensing pool included wearable-based sleep sensor and physical activity sensors. In contrast, none of the physical sensors are theoretically associated with cognitive abilities, so, increasing noise and impairing predictions using social media features which are inherently strong predictors of cognitive ability. An interesting question for future research could establish if sensors that capture speech and communication patterns and social interactions could improve predicting cognitive ability better than social media based predictions alone.

Ease of Interpretation and Domain Adaptations Person-centered analytical approaches can represent individuals on their characteristics rather than defining them simply as a collection of variables [154, 309, 670]. These approaches can be more readily interpretable relative to all-inclusive features, where certain features may obscure others. Rather, this study's approach allows us to cross-introspect features, e.g., how certain offline behaviors are associated with social media language within and across clusters. This kind of explanation and interpretation may be immensely valuable in healthcare and precision medicine, which is defined as *"an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person"* [441]. Such approaches would allow simultaneous inspection at in-conjunction and isolated behaviors. For example, a particular combination of linguistic markers and disrupted physical movements may be useful for early detection of certain mental health symptoms and accordingly guide tailored intervention. Moreover, person-centered approaches can improve human-centered algorithms to be more considerate of the individual in focus by incorporating their circumstances and context.

This study supports Howard and Hoffman's point to researchers of human-centered machine learning that personalized or generalized approaches are not necessarily competitive, but are complementary in terms of methodological, statistical, and theoretical advantages [309]. While one approach may be 'better' outcome-wise, it is worth considering theoretical and practical objectives in understanding the relationships between known and unknown attributes, and in making interventions of addressing a condition. Researchers can benefit from theory-driven computational frameworks that incorporate the predictive capabilities of both precision and generalizability, as well as the explainability in feature interpretation and actionable insights.

Ethical and Privacy Implications

This study has several ethical implications. Proper caution needs to be taken about perceiving or misusing this study as a method to facilitate surveillance or profiling users on user behaviors [461]. It is important to balance the costs and benefits of such systems with an emphasis on privacy-preservation. As Pandit and Lewis describe, "the use of personal data is a double-edged sword that on one side provides benefits through personalisation and user profiling, while the other raises several ethical and moral implications that impede technological progress" [461]. While physical sensor data may be better anonymized and secured compared to demographic information, it is still critical to ensure that the data is simultaneously useful and privacy-preserving.

This study clusters individuals on physical activity data as collected via passive sensors — while these clusters can be characterized on different variables, they may not necessarily translate to mappable individual characteristics, and there is no particular means to simply label them as "desirable" or "undesirable". Any misinterpretation and misuse can bear consequences. For instance, a possible misuse in workplace contexts could be clustering employees on behaviors such as work-times and routines, then characterizing them on productivity, proactivity, and pro-socialness, followed by rewarding and penalizing employees on such characterizations. Any such empirical analyses require careful and in-depth supplemental ethical analyses before enacting any inference and decision-related outcomes.

Finally, with the ubiquity of digital data, the proposed approach of personalizing predictions could be adopted in various contexts, including ones without awareness or consent of individuals. This forms a part of larger discussions on ethical and responsible use of data which require forthcoming discussions among ubiquitous computing researchers, ethicists, and practitioners to understand and respect the individual perspectives on such use of data, which can start from those who choose to participate (or not participate) in multimodal sensing studies [514].

6.2 Understanding Life Event Disclosures on Social Media

Ups and downs are inevitable in our life. As social media platforms continually emerge as important parts of many of our lives [478], they serve many needs and purposes surrounding those very ups and downs of life. Not only do these platforms enable us to connect with others and share day-to-day happenings in life [74, 263, 626], they also have explicit affordances [87] in design that allow us to record and archive their important life events. For instance, the Facebook timeline reminds us of birthdays and personal milestones.

Toward better user experience, most social media platforms today employ algorithms to recommend, rank, or curate personalized content. However, despite providing affordances to gather information on life events, social media content personalization largely relies on topics, interests, and social connections, and rarely accounts for an individual's life events. For this reason, when a Facebook user Eric Meyer was shown his "Year in Review" on the platform in 2014 that included his now-dead daughter's picture, he felt the feature to not only be jarring but also emotionally triggering – labeling the News Feed algorithm as "inadvertently cruel" due to its insensitivity to people's life events [415].

In attempts to serve as safe spaces for authentic expression, support seeking, and promoting wellbeing [91, 165], social media platforms need to consider affordances and algorithms that are sensitive to, respectful of, and compassionate towards major happenings in an individual's life. Such an approach can improve the value one can gain from social media participation, such as meeting varied emotional, informational, and therapeutic needs, and empowering people to gain, maintain, and leverage their social capital. Furthermore, research in human-computer interaction (HCI) and computer-mediated communication (CMC) reveals how naturalistic, self-initiated, and open-ended forms of social data recording, enabled by social media, can augment our understanding of people's reactions and behavior changes surrounding major life events, such as gender transition [284], death of a loved one [83, 404], child birth [164], job loss [91], and pregnancy loss [19]. For example, after a personal crisis, people may desire to reach out to their social media networks for support [19], and following a job loss, an individual may seek empathy from their online social ties and seek new opportunities or job search-related resources from their weak ties [91]. Together, this calls for a critical need to understand social media disclosures of life events.

A life event disclosure on social media uniquely conveys how someone perceives and shares their feelings about the event. However, from an individual's perspective, deciding to self-disclose something as sensitive as a life event on social media can be influenced and compounded by various factors. Literature outlines social media disclosure is affected by factors related to self-presentation, social desirability, audience, boundary regulation, and stigma — people may want to be viewed in particular ways across different audiences, or may not be comfortable about sharing some aspects of their lives with their social media audience [249, 338, 403]. Importantly, an individual may not disclose all life events on social media, and the disclosure choices may vary across individuals and situations. However, the specific factors that explain disclosures (and non-disclosures) remain largely unknown. A deeper examination of life event disclosures would help us understand the authenticity of social media postings regarding how closely this data reflects real-world occurrences of life events in one's life. This would also help to design platform affordances that account for and are sensitive to an individual's life events, and content curation/recommendation algorithms that more adequately represent the gap between observed and unobserved social media behaviors.

Towards designing platforms sensitive to life events, this formative study examines what life events people disclose or withhold on social media, how these disclosures happen, and what are the attributes of individuals who tend to disclose versus not. To accomplish the research goal, in the absence of "true ground-truth" of life event occurrences, I compare social media disclosures of life events with life events self-reported on a standardized survey. Specifically, I use year-long Facebook data from 236 participants who also responded to a retrospective survey, adapted from the PERI life events scale [180], which inquired about life event occurrences in the past year. I ask the below research questions:

RQ 1: How do life event disclosures on social media deviate from a self-reported survey?

RQ 2: How do individual and event attributes explain the deviation in life event disclosure on social media compared to the self-reported survey?

First, targeting the question of *how online self-disclosures of life events deviate from self-reports*, I qualitatively code and define life event disclosures on Facebook data. **This study contributes a comprehensive codebook (available for theory and practice) that enhances our understanding of social media disclosures of life events.** I thematically analyze the language of life event descriptions on social media as compared to their occurrences, with insightful findings such as: social media life event disclosures are typically expressive and emotional in nature; multiple life events may be recorded in the two modalities – social media and survey that might be related, unrelated, or causal; and that negative events tend to stand out in the retrospective recall of individuals, manifested through their survey responses.

Second, given an individual and a life event, I examine *how individual and event attributes explain the deviation in disclosure on social media from self-reported surveys*. I build logistic regression models of logging behaviors by controlling for individual and event attributes. Here individual attributes correspond to demographics and intrinsic traits of cognitive ability, personality, and affect, and event-centric attributes correspond to valence, significance, recency, anticipation, intimacy, scope, status and type of event. The analyses reveal significant findings advancing the understanding of online life event disclosures: positive and anticipated events are more likely to be disclosed online, whereas significant, recent, and intimate events bear a propensity to be self-reported in survey.

The findings reveal how different life events elicit varied decision-making processes on the part of social media users surrounding what, when, and how to disclose, while also navigating the underlying norms of the platform and the audience of a potential disclosure. Then I unpack the fundamental differences between social media platforms and surveys pertaining to the context of use and available affordances, and discuss a need to understand and straddle the socio-technical gap [5] between what individuals disclose online in a self-initiated, intrinsically motivated manner, and what they self-report offline to a prompted survey conducted by a more private but unfamiliar audience of researchers. Drawing on these theoretical underpinnings and implications, this study argues that a "one size fits all" approach to scaffold online life event disclosures may not work. I conclude by providing design suggestions for social computing systems that are sensitive to people's life events, including strategies that accommodate non-disclosure practices and that provide agency to those social media users who choose not to disclose specific life events.

6.2.1 Study and Data

This study uses data from the Tesserae project (section 4.1) [406].

This study examines life event disclosures on the Facebook data of the participants. Given that Facebook is the most popular social media platform [262] and its longitudinal nature has enabled social media studies of individual differences [19, 164], it suits the problem setting of understanding life event disclosures on social media. Among the Tesseae participants, out of the 572 participants who provided access to Facebook data, 242 participants did not make any update during the year-long study period between January 2018 and April

Table 6.7: Descriptive statistics of self-reported demographics and psychological constructs of 236 participants with both social media and life events survey data.

Covariates	Value Type	Values / Distribution
Demographic Characteristics		
Gender	Categorical	Male (n=121) | Female (n=115)
Born in U.S.	Categorical	Yes (n=218) | No (n=18)
Age	Continuous	Range (22:63), Mean = 36.57, Std. = 9.88
Education Level	Ordinal	5 values [HS., College, Grad. Student, Master's, Doctoral]
Cognitive Ability (Shipley scale)		
Fluid (Abstraction)	Continuous	Range (5:24), Mean = 16.93, Std. = 2.94
Crystallized (Vocabulary)	Continuous	Range (22:40), Mean = 33.70, Std. = 3.32
Personality Trait (BFI scale)		
Openness	Continuous	Range (2.17:5), Mean = 3.84, Std. = 0.61
Conscientiousness	Continuous	Range (1.92:5), Mean = 3.94, Std. = 0.63
Extraversion	Continuous	Range (1.67:4.92), Mean = 3.42, Std. = 0.68
Agreeableness	Continuous	Range (2.25:5), Mean = 3.95, Std. = 0.55
Neuroticism	Continuous	Range (1:4.58), Mean = 2.44, Std. = 0.78
Affect and Wellbeing		
Pos. Affect	Continuous	Range (13:49), Mean = 34.24, Std. = 5.69
Neg. Affect	Continuous	Range (10:37), Mean = 16.83, Std. = 4.62
Anxiety	Continuous	Range (20:67), Mean = 37.83, Std. = 9.33
Sleep Quality	Continuous	Range (1:16), Mean = 6.80, Std. = 2.57

2019 — the same period when the participants' self-reported life event occurrences were also collected. This paper uses a data of $14,202$ posts of the remaining 330 participants in subsubsection 6.2.2, which is followed by examining factors for life event disclosures on a subset of 236 participants' data who also responded to self-reported survey on life events, explained below.

Self-Reported Survey Data

As already explained before, Tesserae project included a number of surveys during enrollment, including demographics and individual differences. Table 6.7 summarizes the distribution of the self-reported data within the 236 participants, where we find a reasonably well distribution in demographics and psychological traits among the participants.

End of Participation Period: Life Events Survey Data At the end of the participation period of Tesserae, participants were optionally asked to fill in a life events survey. This life events survey was designed drawing on the Psychiatric Epidemiology Research Interview (PERI) life events scale [181]. Life events were broadly categorized as School, Personal,

Table 6.8: Life Event categories, example hints provided in survey, and example self-reported description in the post-participation self-reported survey — survey scale drawn on the PERI life events scale [181].

Event Type	Category Hint	Example Self-Reported Description
School	Back to school, Changed school, Finished school, Issue at school, etc.	Accepted to business school
Personal	Getting married or divorced, Having a child, Experiencing a death of someone close, Moved residences, Damage of property, etc.	Was working on an adoption
Work	Changed jobs, Received a promotion, Was fired, Had performance review, Received bonus, End of quarter or year, Reorganization	Given more responsibilities in my job, which made me realize I don't want to work in this job anymore
Health	Physical illness or injury, Health treatment, Miscarriage/Stillbirth, Pregnancy related changes, Started menopause, Health changes	Mother diagnosed with kidney failure and congestive heart failure
Financial	Went into debt, Took out mortgage, Made a large purchase, e.g. car or home, Experienced financial gain or loss	Paid off 2 vehicles and refinanced one to pay off high interest credit cards
Local/ Regional	Weather-related changes (blizzard, flood, storm, etc.), Societal changes (political or economic event, sports event, mass-shooting, etc.)	Was at a baseball game where my team advanced to National League Championship
Other	Any other events that do not fall under the above categories	-

Work/Organization, Health, Financial, Local/Regional, and Other. For each category, the survey also included example seed events to help the participant understand respective categories. Participants were briefed that they could refer to their calendars and any relevant personal diaries or journals while completing the survey, to verify the events and dates. The survey was designed in such a way that participants could enter more than one event, and include corresponding attributes about the events. These attributes include a brief description of the event, and two 7-item Likert scales of self-identified significance (Lowest to Highest significance) and valence (Extremely Negative to Extremely Positive) of the life event. In addition, participants entered the start and end date range, status of the event (ongoing or ended), and a confidence value (7-item Likert scale from Lowest to Highest Confidence) regarding the occurrence of the event. Table 6.8 shows the different categories of life events in the survey along with category hint provided in survey and example self-reported descriptions from the responses.

Out of the initial total of 754 participants, 423 participants responded to these surveys with 1,547 entries of life events during the study participation period (mean = 3.86 events per individual). Out of these 423 responded participants, 236 provided us the social media data (above subsection). I examine the data of these 236 participants to understand the deviation of online self-disclosure of life events from self-reports.

6.2.2 Methods

Defining and Annotating Life Event Disclosures on Social Media

Social media facilitates self-disclosures of experiences from day-to-day lives [20, 205]. From the standpoint of life event disclosures, social media posts are unstructured forms of textual expressions, and this data lacks "ground-truth" labels regarding what constitutes a life event disclosure and what does not. So I first aim to systematically identify online self-disclosures of life events from social media data with respect to a theoretical grounding of life event occurrences. I adopt a qualitative coding approach to iteratively define and annotate life event expressions on social media. This study primarily builds on and adapts the list of categories from the PERI life event scale [181] in the context of social media data. This theory-driven coding enables us to formally define a social media post to contain a life event disclosure *if the post describes an event which is directly or indirectly associated with the individual or their close ones, such that it potentially leaves a psychological, physiological, or behavioral impact, or be significant enough to be remembered after a period.*

While the PERI life events scale [180] identified a list of various life event categories, there is no established means to adopt this on social media data. This study applies these categories in a sort of directed coding approach [310], i.e., when developing the codebook, this study allowed concepts and meanings to emerge from posts in somewhat of an open coding [602]. This codebook is particularly driven towards identifying life event disclosures from social media language. The Appendix provides the detailed codebook (Table C.2) to identify life event disclosures on social media.

This study recruited five annotators who are undergraduate students. Although the Facebook data primarily consists of English posts and belongs to a participant pool recruited in the U.S., all participants were demographically and culturally heterogeneous. Therefore, it is important to note that the annotators (three women and two men) belonged to diverse cultural backgrounds; in race/ethnicity, two identified as Caucasian, two as East Asian, and one as South Asian. During discussions, the research team found specific occurrences when annotators were able to identify culturally significant events due to their cultural backgrounds, which could have been missed by other annotators. These five annotators first coded a random sample of 140 Facebook posts with the PERI life events scale [180] and the instruction that they could add new categories if a post was a life event disclosure and it did not fit any of the existing PERI categories.

The annotators and the research team then discussed the coding one by one in detail. Together, decisions were made on all posts with coding discrepancies, and the codebook was revised based on agreeable themes. These included resolving boundary and similar sounding cases such as identifying a *trip* versus a *vacation*. Next, the annotators separately coded an additional 50 randomly selected posts. For the total 200 posts, the research team found a high agreement of 88% between the annotators and an average Fleiss κ of 0.71. Two annotators then independently coded the remaining 14,002 posts. Because of the subjectivity in social media data, I adopted a liberal identification strategy that a post is labeled as a self-disclosure of life event if it is labeled so by either of the two annotators. The research team discussed several explicit and boundary cases to decide general criteria for identifying life event disclosures, which I elaborate on in Table 6.9. Note that the presence of a post within the context of other posts (before and after it) drove the decision-making towards labeling a post.

Comparing Life Events Disclosed on Social Media Versus Reported on Survey

So far, I described our approach to obtain life events disclosed on social media and self-reported in surveys. Consequently, for the common set of 236 participants for whom I have both modalities of data, I obtain 912 life events self-reported on the survey and 1,669 self-disclosed on social media. To answer the core research question on *what, how, when,*

Table 6.9: Brief strategies and considerations to identify life events on social media.

What constitutes a life event disclosure?
Present events with potentially significant impact in the future.
Posts were coded as life events disclosing a present event which is significant enough that to be recalled in a few years, or if the event in disclosure could potentially leave a significant emotional impact in the future. For example, *“Horrible day for travel. Two canceled flights and 2 delays. Sharing the sights from this week while I wait to get home.”*
Past events with significant emotional impact in the present.
Self-disclosures about recalling events from the past. This conveys the significance of the event in the individual’s life and leaves emotional impact. Therefore, for events that occurred a while ago, if they have a big enough emotional impact even in the present, these posts would be identified as a life event, e.g., recollecting the death of someone close often results in grief in the present [466], such as, *“When you are looking for one child’s birth certificate and find the other child’s death certificate.. 33 days and you would be 16..”*
Using the post wording.
Wherever applicable, in cases of close tie in assigning a post with a life event category, we prioritized the wording in the post. We considered that the individual’s self-description of an event is less biased and closer to self-perceived life event type. For example, when deciding between *trip* and *vacation*, if the post explicitly used either of the two words, we assigned the same life event category. For example, we assigned *trip* for *“For my recent business trip I flew Delta. I’m giving them 4 stars. They have on-demand in-flight movies and I got to watch Black Panther.”*
Underlying reason of an event.
As above, when multiple categories could fit a post, we prioritized the one that seemed to be the underlying cause. Sometimes, other posts around the same date provided more context to make these decisions. For example, in the following post, although both *vacation* and *positive relationship* could be appropriate, *positive relationship* (anniversary) was the more underlying cause (also consistent with the individual’s other posts around the same date), *“What a beautiful weekend celebrating our 10th Anniversary! So thankful for getting away to enjoy time together as husband and wife <3.”*
Disclosing multiple life events.
A post may disclose multiple life events, including continuous or ongoing events, e.g., a vacation may include a birthday party, or a wedding planning post may also talk about different investments, e.g., *“Going to start selling a small selection of simple car [..] Trying to make some money on the side for wedding and honeymoon, and my medications. Also gotta pay this damn hospital bill now.”*
Continuous Life Events
Certain posts may mention continuous life events. The longitudinal data enables identifying events lasting for a time period, e.g., *start*, *during*, and *end* of a vacation. Continuous events can be 1) a series of posts which together build a continuous event, 2) other posts providing context about a seemingly vague post at hand, and 3) a single post describing a continuous event. These may not be exclusive and can co-occur, e.g., a post describing a *“view”* or a *“beautiful city”* may seem vague, but, posts around the date provided context these are during-vacation activities. Again, a continuous life event can include related or unrelated life events within that period.
Additional Life Events Categories
While annotating social media life event disclosures, we included open coding, allowing new categories, which might not directly be present in the PERI scale. For example, we added a new category of *Voted* for a post, *“I voted”*.
What does not constitute a life event disclosure?
Vague Post.
Exclude if the posts is too vague to make a deduction of a life event, e.g., *“Waited for this FOR FOREVER!!!!!!!!!.”*
Joke or Entertainment Media related.
Found posts that mention a life event, or keywords related to life events, which were explicit expressions of these to be a joke, or a description about an event in a movie, TV show, video game etc, for example, *“The end... he died lol!”*
Past events, but no significant emotional impact in the present.
Posts describing events or self-experiences from the past, without significantly affecting the present, were excluded, e.g., *“The meals, and especially the Blue Mountain coffee, were the best in Jamaica.”*
General shares or global events.
Posts in third-person of generic information (no personal reference) based sharing were excluded, e.g., *“In four years as a student at University, Name had seven internships.[..] The experiences helped her decide what she wants in a career [..]”*

and by whom life events are disclosed on social media compared to self-reported surveys, first, I examine the distribution of life events in the two modalities of datasets. Then, I conduct a thematic analysis of the overlapping life event logs from the two datasets. Finally, I examine the factors that explain the overlap and deviation in reportage on either or both the modalities, for which, I describe the statistical tests in the following subsection.

Examining Factors Associated with Life Events Disclosures and Survey Self-Reports

To examine the factors explaining deviation in recording life events on the two modalities, I identify a set of theory-driven covariates that may contribute to an individual's life event disclosure (or no disclosure) on either or both the modalities. I use these covariates in our statistical tests and models to explain such life event disclosure.

Covariates Given an individual and a life event, the covariates belong to two major kinds — *individual centric attributes* and *event-centric attributes*, which I describe below.

Individual-centric Attributes Given that an individual's disclosure is known to be driven by their demographic and intrinsic traits, I use individuals' demographic and psychological attributes (as in Table 6.7) in our models.

Demographics. Prior studies controlled on several demographic attributes in studying self-presentation and self-disclosure of individuals [536]. I include demographic variables of gender, age, born in the U.S., educational level, and income in our models.

Cognitive Ability. Cognitive ability is known to associate with an individual's disclosure and expressiveness [493], which I include as independent variables in our model. I used the the Shipley scales of 1) Abstraction measuring fluid cognitive ability and 2) Vocabulary measuring crystallized cognitive ability [578].

Personality. Prior work revealed the role of personality in people's disclosure, including in online settings [308, 561]. I include personality trait as a covariate in our models where ground-truth assessments of personality traits come from the Big-Five inventory along the traits of openness, conscientiousness, extraversion, agreeableness, and neuroticism [589].

Affect and Wellbeing. Social media use is known to be associated with people's trait based measures of affect and wellbeing [655]. I include positive and negative affect traits as assessed by the PANAS-X scale [658], anxiety trait as assessed by the STAI-Tait scale [591], and sleep quality as assessed by the PSQI scale [182]. I note that PSQI scale assesses sleep quality in such a way that lower values indicate healthier sleep. Therefore, for easier interpretation, I reverse-scale the values and use "Healthy Sleep Quality" as a covariate which directly correlates with healthier sleep.

Event-centric Attributes People's life event disclosures (or non-disclosures) may be driven by event-centric attributes. I describe the motivation and the operationalization of event-centric attributes considered in our models below.

Event Recency. Self-reported surveys are known to be biased to more recent events [29, 246]. However, no such evidence exists about social media postings, which is more of a self-initiated and in-the-present recording. To understand such an effect in online life event disclosure, I include recency of events as an independent variable. I first choose a reference date as the date of conducting the end of participation survey. Then, for the survey data, I calculate the number of days between the reference date and the self-reported occurrence of event (also collected in the survey). For social media data, I calculate the number of days between the reference date and the date of posting. For easier interpretation and standardization, I reverse-scale the number of dates to obtain recency on a min-max scale of 0 to 1 — such that 1 represents most recent event whereas 0 represents least recent events.

Event Significance. Individuals are known to be more likely to recall and report events which bear greater degree of significance in their lives in whatsoever ways [436]. This aligns with survival salience [436], and emotional or informational relevance can drive the salience in memory [344, 470]. Participants self-reported how significant they considered each life event they logged — which I use as an independent variable for event records from surveys. For events recorded on social media, I adopt the significance rating per event as per the PERI life events scale [181]. I separately standardize the significance scores on a min-max scale of 0-1 to make the significance scores comparable across the modalities, and then use this scaled score as an independent variable in the models.

Valence. The independent variables include valence or sentiment of the event, in terms of being positive or negative. Like above, valence directly associates with emotional relevance of an event in the memory [344]. The survey data included people's self-reported valence on a Likert scale of extremely positive to extremely negative, which I group into three bins of positive, neutral, and negative to minimize subjectivity in the analyses. To score valence of social media life events, I use the VADER tool [315] to identify the major sentiment of a post among positive, negative, and neutral, which I use as the valence for life event entries from social media data.

Anticipation of an Event. Life events include a characteristic on the basis of anticipation: Compas, Davis, and Forsythe defined anticipated events as the events which an individual can either hope or worry about in the next six months [132]. I adopt a similar definition to label each life event in the dataset with binary labels of anticipated or unanticipated. Example anticipated events are buying a house, childbirth/pregnancy related events, whereas example unanticipated events are accidents or getting fired from work.

Intimacy in Disclosure. Prior work studied that intimacy is a core attribute that might moderate people's disclosure behavior [15, 206, 389]. Intimacy relates to the degree to which one can comfortably open up about a particular event at personal, close, trusted others, and public circles of relationships [206, 249]. While social media disclosures are broadcasted to some form of public or known private audience, a self-reported survey is likely self-perceived to be much more private. I draw upon the annotation scheme from Ernala, Labetoulle, Bane, Birnbaum, Rizvi, Kane, and De Choudhury to code life event descriptions — I annotate both survey self-reports and online disclosures of life events on a degree of intimacy Likert scale of Low, Medium, and High[2].

Scope of an Event. The social ecological model posits that an individual's wellbeing is impacted by different layers of scope ranging across individual, relationships with close ones, societal, and local factors [102]. Similarly, the scope of a life event can either be

[2]The Appendix provides the detailed codebook, and the codebooks to annotate intimacy, scope, and status.

directly on the individual themselves, or their close ones, or something more generic [59]. I label each life event in the datasets with their ecological scope of directness on a three-point Likert scale[1] such that 1) *Low* scope events include generic events such as bad weather or neighborhood related events, 2) *Medium* scope events are associated with someone close and leave an indirect effect on the individual (e.g., spouse's pregnancy, child going to school), and 3) *High* scope events are unique to and direct on the individual, e.g., being fired from job themselves.

Temporal Status. I also include temporal status of events in terms of a binary value of *ongoing* or *ended*. This factor takes into account during-reporting continuity of events. Our survey included self-reported entries of the status of event, and for social media, I manually identified the temporal status by going through the life event disclosure posts[1].

Event Type. As introduced earlier, the datasets (both social media and surveys) group the life events into six broad categories of *School, Health, Personal, Financial, Work, and Local*. While the self-reported survey data was already annotated with these categories by the participants, the social media data life event expressions were annotated by the annotation approach and codebook[1]. I use the categorical variable of life event type as covariates in the analyses. Besides, although the data contains labels of finer categories of life events (e.g., *vacation, health loss, bad weather, child birth*, etc.), the number of records per event is plausibly not significant for statistical power, and may lead to inconclusive or misleading results [129] In addition, theoretically an individual only experiences a limited number of life events per year [251, 405], so it would be impractical to include all possible life events without a significantly larger sample size than what I have. I validate this hypothesis by conducting a χ^2-square test, which reveals $\chi^2 = \texttt{NaN}$ and $\texttt{p} = \texttt{NA}$, suggesting not enough observations per finer categories of life event.

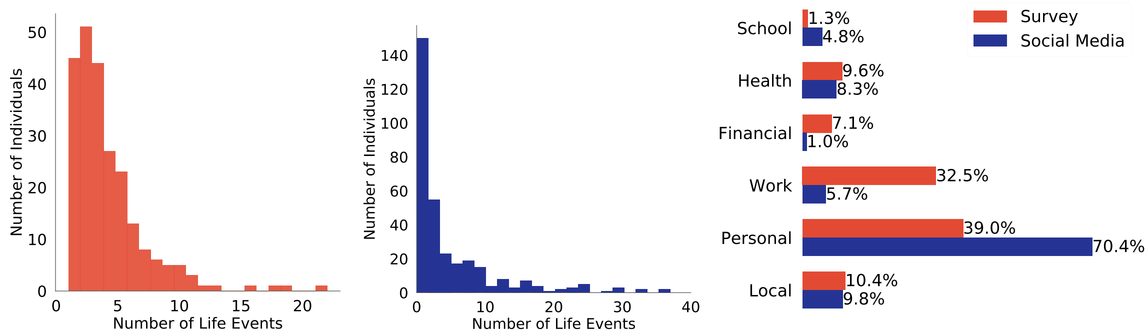
Baseline Attributes Social media and self-reported surveys are fundamentally two different data modalities, and it is important to control the models on an individual's baseline behavior on these modalities. Essentially, for each individual, I compute four baseline attributes — *social media baseline attributes* include, 1) total number of posts and 2) average length of post per individual, and *survey baseline attributes* include, 3) total number of responses and 4) average significance self-reported in each response. These baseline attributes go in as covariates in the models.

Tests and Models First, I obtain a union of all the life events recorded on social media and on survey as the total dataset ($\mathtt{D_T}$). Then, I conduct a One-way Multivariate Analysis of Variance (MANOVA) tests on the combination of dependent variables of social media self-disclosure and survey self-report to the set of theory-driven covariates explained above. A statistical significance in MANOVA would reveal the importance of each covariate in explaining life event reportage on either or both of social media and surveys.

Next, to understand the direction of the factors in their associate with life event disclosure, I conduct two kinds of analyses drawn on nested logistic regression models — one on $\mathtt{D_T}$ and the other on a subset, $\mathtt{D_S}$ consisting of events recorded in one of the two modalities. This would allow us to examine the intricacies of each factor and their signed (positive or negative) importance in explaining reportage. I describe the two analyses below:

- *Convergence:* The first analysis studies whether a life event is likely to be recorded in *both social media and survey* modalities. On $\mathtt{D_T}$, I build a binary logistic regression model that uses dependent variable as a binarized value based on the occurrence on both modalities, i.e., if the event is logged on both modalities, it is labeled as 1, otherwise 0. This model is referred to as $\mathbf{Model_1}$.

- *Divergence:* The second analysis is conducted on $\mathtt{D_S}$, among life event records which are *not* recorded on both the modalities — what is the likelihood of it to be self-disclosed on social media versus self-reported on survey. This logistic regression model uses as dependent variable the binarized value based on the occurrence on either of the modalities. That is, given an individual's life event log which does not



(a) Survey Self-reports (b) Social Media Disclosures (c) Category-wise Distribution

Figure 6.5: Distribution of data by life events (a) in self-reported survey data, (b) in social media data, (c) per category (percentage values are on all the life events reported within each dataset).

occur at both modalities, it is labeled as 1 when self-disclosed online, and labeled as 0 when self-reported on survey. I refer this model as $\mathbf{Model_2}$.

6.2.3 Results

Distributions of Life Events

Table 6.10: Top life event recorded in survey self-reports and social media self-disclosures.

Survey Self-reports		
Life Event	**Type**	**Count**
Vacation	Personal	182
Performance Review	Work	117
Bad Weather	Local	88
Health Loss	Health	88
Promoted	Work	53
Positive Job Switch	Work	45
Heavy Work	Work	44
Got Bonus	Work	44
Neutral Job Switch	Work	42
Trip	Personal	40
Installment Purchase	Financial	36
Child Birth	Personal	33
Death in Family	Personal	28
Positive Move	Personal	28
Financial Gain/Loss	Financial	27

Social Media Self-disclosures		
Life Event	**Type**	**Count**
Vacation	Personal	485
Trip	Personal	227
Increased Social Activity	Personal	142
Family Meetups	Personal	106
Positive Relationship	Personal	85
Health Gain	Health	69
New Hobby	Personal	67
Positive Move	Personal	56
Death in Family	Personal	45
Back to School	School	42
Work Success	Work	40
Remodeled Home	Personal	34
Good Worklife	Work	34
Injury	Health	34
Child Birth	Health	29

I present the distribution of life events reportage on both modalities by number of individuals in Figure 6.5a and Figure 6.5b. First, there is a heavy skew at x=0 for social media disclosures, which does not exist for survey self-reports — a key difference in the characteristic of the two data modalities. Out of the $14,359$ Facebook posts, only 14%

$(2,031)$ express life events as per the annotation. In contrast, the survey is a dedicated effort directly asking the participants to log life events, so, 100% of its responses correspond to some form of self-perceived notion of life event per individual.

Next, Figure 6.5c shows the category-wise distribution of life events in the two modalities. Both the datasets show a prevalence of *Personal* life events — 39.5% among all survey responses, and a high 70.4% among all online disclosures. Interestingly, *Work*, which is significantly logged in survey self-reports (32.5%), appears low on social media (5.7%). *Health* events are recorded comparably on both surveys (9.5%) and on social media (8.3%).

Table 6.10 presents the top life events recorded on the two modalities. *Vacation* scores the highest on both. In fact, *vacations* and *trips* occur more commonly across individuals as opposed to the rarity and uniqueness of other events. Our data suggests that Facebook's design and perceived use-case may facilitate individuals to post prevalently about *vacation* and *trip* events. Again, these events are often recorded on calendars, which may guide individuals to report these events in the post-participation life events survey.

Table 6.10 also explains the significant occurrence of other categories in the self-reported survey data including, *Work*-related *performance review, promotions, heavy work*, and *job switches*, none of which are disclosed significantly on social media. Rather, the only *Work* categories frequently disclosed online are *good worklife* and *work success* — both of which bear a positivity in valence. This may indicate that people are not comfortable about sharing work-related negativity on social media due to concerns of employer surveillance [236]. Another interesting contrast includes that *health loss* appears as a top event self-reported in surveys, whereas *health gain* occurs in those disclosed online. These observations suggest an inclination towards disclosing positive events on social media, which may associate with perceived self-presentation and social desirability of individuals on a public platform (social media) [305].

There is difference in labeling life events in the two modalities (self-perceived vs. inferred). This distinction may indirectly explain the observation that the annotation scheme identified increased social activities (e.g., celebrations, gatherings) as "life events", which might not be self-perceived the same way to be recalled during a survey that happened after a period of time. In contrast, *death in family* and *child birth* commonly occur in the top life events on both modalities. These events are known to bear both short-term as well as long-term effect on one individual's life [180].

Language of Life Event Disclosures

I investigate the relationships between social media posts that were temporally similar to life events self-reported on surveys. In particular, for each individual, I inspect events that were overlapping on the two modalities or occurred less than 7 days from each other. I aim to qualitatively determine what relationship, if any, there is between the reportage of life events on these modalities. After identifying pairs of potentially overlapping events from each modality, I compare and code the similarities and differences in linguistic descriptions of the events from the two modalities. Then, based on the codes, I conduct a thematic analysis to gradually coalesce the codes into themes of associating online disclosures and survey reported life event descriptions. Some notable themes are listed below.

Emotional and Expressive Content Social media posts are more likely to bear an emotional tone about events. These include several occurrences for events such as adoption of pet and child birth, *"Name was born today. She was 8lbs 5oz and 21 inches long. We love her so much and are very thankful that she is happy and healthy! Thanks for all of the prayers!!"*. Similarly, social media posts also contain greater and richer detail about the event, for example, someone whose self-report survey entry only recorded a vacation, had posted on their social media about their vacation and positive relationship event, *"Best date night with my husband! Love you to the moon and back dear husband #wefishtogether."*

Co-occurring and Related Events. Sometimes the social media post can reflect a co-occurring and related event in someone's life. For example, an individual who self-reported on the survey to be on a vacation on certain dates, posted about a family meetup during those dates, *"Had the joy and privilege of seeing my niece dance in the ballet Sleeping Beauty today...also got to spend time with some people dear to my heart."*, here vacation and family meetup co-occur. Another individual, who changed jobs, posted about their move to a new city, *"Just rolled into California. Quite some driving but an easy roll into SF tomorrow."*

Followup or Cause-Effect Related Events There are instances where one life event may have triggered or caused a separate life event about which the individual posted on social media. For example, an individual who reported to be assaulted on a particular day, followed up with a Facebook post on *"I'm moving."* Again, an individual who self-reported about a bereavement leave at workplace on survey, had self-disclosed about the death of a family member a day prior to the reported date, *"This guy will be missed. Wish we had more time together [..]."*

Co-occurring but Likely Unrelated Events Interestingly, there are instances of events that co-occur but are likely unrelated to each other. For example, an individual who self-reported on the survey having trouble with their boss at workplace, self-disclosed about their pet on social media, *"Help me find my foster pup a forever home! He is the sweetest and needs a great home asap [..]"* Again, another individual who self-reported on the survey about the death of a pet, had posted about a family reunion during the same time on social media, *"A family reunion time."*

Negative Stands Out in Recall. I find instances where a negative event within a span of events outweighs the rest, and it is the only event reported in the survey (which happens later). In contrast, the social media data archives events from the past but were presumably recorded in-the-present. For example, in one instance, an individual posted about their ongoing vacation on social media, however, in the survey they only logged about a breakup on those dates. On another instance, an individual's social media data revealed them enjoying

Table 6.11: Multi-variate Analysis of Variance (MANOVA) results, * p<0.05, ** p<0.01, *** p<0.001.

Demographic/Trait	Pillai	F	p
Age	0.036	47.01	***
Gender	0.072	100.34	***
Born in US	0.004	4.87	**
Education	0.051	16.73	***
Shipley: Abstraction	0.057	76.61	***
Shipley: Vocabulary	0.033	42.97	***
Personality: Openness	0.002	2.56	*
Personality: Conscientiousness	0.022	28.21	***
Personality: Extraversion	0.003	4.21	*
Personality: Agreeableness	0.077	106.63	***
Personality: Neuroticism	0.030	39.64	***
Positive Affect	0.003	3.73	*
Negative Affect	0.005	7.03	***
STAI: Anxiety	0.001	1.56	
PSQI: Healthy Sleep Quality	0.011	14.45	**

Event Attribute	Pillai	F	p
Valence	0.063	85.43	***
Significance	0.317	592.16	***
Recency	0.044	59.81	***
Anticipation	0.086	120.23	***
Intimacy	0.002	3.11	*
Scope	0.005	1.67	*
Status	0.516	988.62	***
Type	0.183	51.31	***

Baseline Attribute	Pillai	F	p
SM: Num. Posts	0.088	121.00	***
SM: Avg. Post Length	0.003	4.04	*
SR: Num. Records	0.108	152.10	***
SR: Avg. Significance	0.019	25.47	***

a vacation with friends, however they only self-reported a car-crash that might have happened then.

Factors Explaining Life Event Reportage

Importance of Covariates in Reportage First, I examine the importance of individual-centric and event-centric covariates to understand life event disclosures. I conduct MANOVA tests with respect to the Pillai-Bartlett trace, which is considered to be robust and not strongly linked to normality assumptions the data distribution [449]. Table 6.11 summarizes the MANOVA statistics, where the F-statistic quantifies the association of the covariate with the dependent variables, and larger values indicate greater statistical importance. I compare the F-statistic and significance across the covariates. Among the individual attributes, agreeableness (F=106.63) shows the greatest association, closely followed by gender (F=100.34). Among event attributes, status (F=988.62) and significance (F=592.16) show the greatest association, followed by anticipation (F=120.23) and valence (F= 85.43). The statistical significance shown by all variables (except anxiety) empirically validates the choice of the considered theory-driven variables.

Table 6.12: Model_1 (Convergence): Coefficients of linear regression of relevant covariates as independent variables and reporting on both modalities as dependent variable, * p<0.05, ** p<0.01, *** p<0.001. Bar length is proportional to the magnitude of coefficient; for significant rows, **orange bars (positive coefficients)** indicate a propensity to record on **both social media and survey**, whereas **teal bars (negative coefficients)** bars indicate a propensity to record on **one of the modalities**.

Demographic/Trait	Std. Coeff.	p
Age	−0.03	**
Gender: Male	−0.41	***
Born in US: Yes	0.41	
Education: H. School	1.57	***
Education: College	1.33	**
Education: Grad School	1.78	**
Education: Doctoral	1.31	*
Shipley: Abstraction	−0.03	
Shipley: Vocabulary	0.05	**
Personality: Openness	−0.24	
Personality: Conscientiousness	−0.25	*
Personality: Extraversion	0.04	
Personality: Agreeableness	0.49	***
Personality: Neuroticism	0.06	
Positive Affect	−0.04	*
Negative Affect	0.06	***
Stai: Anxiety	−0.03	*
PSQI: Healthy Sleep Quality	0.02	

Event Attribute	Std. Coeff.	p
Valence: Positive	0.24	
Significance	−0.33	***
Recency	−0.24	
Ancptn.: Anticipated	0.16	*
Intimacy	0.08	
Scope	−0.51	**
Status: Ongoing	1.08	***
Type: Health	0.82	**
Type: School	0.54	*
Type: Work	−0.61	*
Type: Local	−0.60	**
Type: Financial	−0.49	**

Baseline Attribute	Std. Coeff.	p
SM: Num. Posts	0.48	***
SM: Avg. Post Length	0.50	
SR: Num. Records	0.33	**
SR: Avg. Significance	0.20	**

AIC = 2047.40, Deg. Freedom= 33, Log-lk. = −988.71, χ^2= 408.98, McFadden's Pseudo R^2 = 0.18, p<0.001

Convergence: Reportage of Events on Both Social Media and Survey $\texttt{Model}_1$ examines the factors associated with life events reportage on *both of* against on *one of* the modalities (ref: Table 6.12). $\texttt{Model}_1$ shows a McFadden's pseudo R^2=0.18, $\chi^2(34)$=408.98 and $p < 0.001$, suggesting that the model is significantly better than an empty model. For a covariate x showing a standardized coefficient estimate of e with statistical significance, could be interpreted as: a change in one unit of standard deviation likely results in e standard deviation change in the log odds of the dependent variable. In the case of $\texttt{Model}_1$, a positive coefficient indicates a propensity to reporting a life event on both modalities, and a negative coefficient indicates a propensity to report on one of the modalities.

Among demographics, the findings suggest that the likelihood to report on both modalities lowers as age increases. Similarly, males are less likely to report on both. This aligns with prior work [44] that males tend to self-disclose lesser than females on certain personal life events. Among traits, crystallized cognitive ability shows a significant positive coefficient. This is plausibly related to the notion that greater cognitive ability is known to drive the ability to distinguish positivity and negativity of situations to accordingly structure emotional expressiveness [510]. In personality traits, conscientiousness and agreeableness are significant, each showing opposite association — conscientiousness negatively associates whereas agreeableness positively associates with the likelihood to report on both modalities. Conscientiousness characterizes one's thoroughness [589] — a significance may be associated with individuals being methodical in delineating what they want to disclose on social media. On the other hand, agreeableness characterizes warmth and friendliness — an individual scoring high on agreeableness likely "gets along well" with others [589, 617]. This plausibly relates to people knowing their online audience better, and experiencing low inhibition to report on both modalities. Affect and wellbeing traits show weak relationships, and interestingly positive and negative affect exhibit opposite directions — higher positive affect explains lower reportage, whereas higher negative affect explains greater reportage on both modalities.

Among event attributes, significance bears a strong negative coefficient ($\mathtt{e}$=-0.33) indicating that significant events are less likely to be reported on both modalities. Anticipated events are likely to be reported on both ($\mathtt{e}$=0.16); these events bear some form of planning or apriori awareness (e.g., child birth), and people may not only disclose them online, but also recall and report them in retrospective survey. In contrast, unanticipated events plausibly relate to emergency circumstances, and people may deprioritize an immediate online disclosure. These could also be short-term events (e.g., a positive relationship act) which may be disclosed on social media in-the-present, but may not remain in one's long-term memory to be self-reported in a survey which happened after a while. Among event types, *Health* and *School* events have propensity to be recorded on both social media and surveys, whereas, *Work* and *Financial* events are unlikely to be recorded on both modalities.

Finally, note the the statistical significance of controlling for baseline behavior of individuals. Recording on both modalities shows a positive association with individuals who typically have more social media posts, more survey records, and whose baseline average

Table 6.13: Model_2 (Divergence): Coefficients of linear regression of relevant covariates as independent variables and reporting on either modality (1 for online/social media and 0 for survey) as dependent variable, * p<0.05, ** p<0.01, *** p<0.001. Bar length is proportional to the magnitude of coefficient; for significant rows, **blue bars (positive coefficient)** indicate a propensity to record **only on social media**, whereas **red bars (negative coefficient)** indicate a propensity to record **only on survey**.

Demographic/Trait	Std. Coeff.	p	Event Attribute	Std. Coeff.	p
Age	0.04	***	Valence: Positive	0.45	***
Gender: Male	−0.38	*	Significance	−1.40	***
Born in US: Yes	−0.75		Recency	−3.56	***
Education: H. School	0.43		Anticipated	0.45	*
Education: College	0.43		Intimacy	−0.75	**
Education: Grad School	0.47	*	Scope	−0.93	***
Education: Doctoral	0.48		Status: Ongoing	3.62	***
Shipley: Abstraction	−0.12	***	Type: Health	−0.98	
Shipley: Vocabulary	−0.05	*	Type: School	0.18	
Personality: Openness	0.18		Type: Work	−1.18	***
Personality: Conscientiousness	−0.04		Type: Local	−1.11	*
Personality: Extraversion	0.13	*	Type: Financial	−2.90	***
Personality: Agreeableness	0.73	***			
Personality: Neuroticism	−0.11		**Baseline Attribute**	**Std. Coeff.**	**p**
Positive Affect	0.03		SM: Num. Posts	0.90	***
Negative Affect	−0.05	*	SM: Avg. Posts Length	−1.59	
Stai: Anxiety	0.04		SR: Num. Records	0.49	***
PSQI: Healthy Sleep Quality	−0.04		SR: Avg. Significance	−1.57	***
AIC $= 628.26$, Deg. Freedom$= 34$, Log-lk. $= -279.13$, $\chi^2= 1785.83$, McFadden's Pseudo $R^2 = 0.77$, p<0.001					

significance of self-reported life events on survey is higher. However, average length of social media posts shows no statistical significance with respect to recording behavior.

Divergence: Reportage of Events on Social Media Versus on Survey $\texttt{Model}_2$ examines the factors that associate with reporting life events on *either of* the two modalities (ref: Table 6.13). $\texttt{Model}_2$ shows a McFadden's pseudo R^2=0.77, χ^2(34)=1785.83 with p<0.001, i.e., the model is significantly better than an empty model. Here, positive coefficients suggests a propensity to record online, and negative suggests a propensity to report on survey (and not social media).

Among individual-centric attributes, males (e=−0.38) show a lower likelihood to self-disclose online. This observation somewhat supports prior work that found men to show lower online self-disclosure than women [575]. There is a strong association with agreeableness (e=0.73) — indicating that individuals with greater agreeableness have a likelihood to self-disclose life events on social media. Similarly, extraversion shows a positive coefficient

(e=0.13). Extraversion characterizes one's outgoing, talkative, and energetic behavior [617], and this trait is known to associate with greater expressiveness and disclosure [453, 507]. There is a weak negative significance for negative affect (e=-0.05), indicating that individuals scoring high on negative affect are less likely to disclose on social media, which could be associated with privacy and audience perceptions as noted in prior work [151, 403].

Among event-centric attributes, valence (e=0.45) and anticipation (e=0.45) bear positive coefficients. This suggests that individuals tend to mostly disclose events on social media that are positive and/or that are anticipated. On the other hand, both significance (e=-1.40) and recency (e=-2.90) bear strong negative coefficients. This supports prior research regarding the bias of self-reported surveys due to retrospective recall and significance of events [620]. Also, intimacy (e=-0.75) and scope (e=-1.03) bear negative coefficients, likely related to the public-facing nature of social media and people's self-presentation. Unsurprisingly, social media disclosures are also more skewed towards ongoing events because they enable in-the-present sharing, unlike surveys that elicit retrospective recollection.

Among life event types, *Financial, Work,* and *Local* events bear low likelihood to be disclosed on social media. People may not be comfortable about sharing their financial gain or loss events publicly on social media, or they may not share work-related events, especially if they have concerns around context collapse [403]. In contrast, *School* events may not be deemed that private, and people may be comfortable sharing about school-related success and milestones.

Finally, among baseline attributes, number of social media posts positively associates with life event disclosures on social media. Again, number of survey records also positively associates with social media disclosure. However, individuals who reported higher significance of events on average tend to post lower on social media — this could relate with people's baseline perceptions of event significance and social media disclosures.

6.2.4 Discussion

Theoretical Underpinnings and Implications

We sought to examine how/when people tend to disclose life events on social media, and the attributes of individuals who choose to disclose versus not. This section first discusses the theoretical underpinnings of this study, drawing on a host of theories and conceptualizations in social computing and HCI. To situate the validity of the disparities between disclosure and non-disclosure, we compared and contrasted social media disclosures with survey self-reports of life events — the latter being the gold standard in capturing life events. Accordingly, we also discuss how some of these differences are rooted in the differences in the two modalities in their context of use and available affordances.

The Role of Audience and Norms Social desirability is a known bias in surveys and in face-to-face offline settings [265]. This study reinforces prior evidence that this factor could potentially modulate social media disclosures as well [430]. We found instances when individuals were comfortable to disclose positive or anticipated events on social media that were not remembered during the survey. This may indicate a varied set of self-presentation goals propelled by the positivity bias in normative Facebook use [90], or a desire for selective "performance" as per Goffman's "frontstage/ backstage" metaphor for impression management and social roles enactment [249], or for exhibitionism [305], or for receiving instant or short-term social approval and gratification [662].

These disclosures may also stem from a need to maintain and bridge social capital around transitory or minor happenings in one's life, where sharing certain milestones, such an imminent wedding, or leaving/starting a job has become customary — a recent survey found that people "prefer sharing life's milestones with their social network than in person" [80]. In fact, sharing life milestones on social media may not only revive dormant social connections, and simultaneously elicit responses or communication from an individual's passive or weak ties [594], but also enhance the emotional tone and impact of the event [123]. Finally, positive and anticipated life event disclosures may also be attributed to the "desire to use online social media as a way for archiving life experiences and reflecting on identities," especially if the events are associated with liminality [260]. Taken together, the findings shine a light on how the underlying norms of a social media platform, as well as its relationship to social desirability and impression management, may impact the semantics of a life event from an individual's perspective, and the decision surrounding its online disclosure.

Complementarily, as societal norms motivate people to behave in particular ways [550], a social media platform's norms may encourage certain disclosures as well as impose certain expectations that discourage people from sharing specific life events. Drawing on the literature on social comparison in social media [90, 462], people may not disclose very sensitive events such as an extra-marital relationship, a family conflict, or pregnancy loss for fear of social disenfranchisement, stigma, or shame [18]. This study found that individuals withheld disclosing work-related and finance-related events on social media despite their occurrences per self-reports on the survey. Building upon the Disclosure-Decision Making framework proposed by Andalibi [18], one can conjecture that these decisions may be driven by people's specific imagined or actual audiences [403] including their mental representations [437], wherein, due to concerns of context collapse [403], conflicting social spheres [64], surveillance [236], or the (semi-) public nature of the platforms [305], certain life events may be deemed less appropriate or share-worthy compared to others. Moreover, this study found a lower likelihood of disclosing particularly intimate events or events too personal in their scope on social media. The design of the Facebook platform may in itself be a key factor driving self-regulatory decisions of non-disclosures [173]. Facebook particularly does not enable anonymity, a factor known to be facilitating intimate content sharing [389]. With an emphasis on "integrity and authenticity" as a community standard on the platform[3], other known disclosure risk mitigation strategies such as switching communication channels [271], using multiple accounts [645], or sharing incorrect information [340], may not apply for life

[3]https://www.facebook.com/communitystandards/integrity_authenticity

event disclosures on Facebook.

Summarily, I draw upon Newman, Lauterbach, Munson, Resnick, and Morris's [437] observations about sharing sensitive information on Facebook, that people carefully navigate the tension between sharing vulnerability, needs, and health status information and the desire to convey positive images of themselves. The findings point out an apparent dichotomy that the same factors which encourage disclosure on Facebook (e.g., real identity, online and offline friendship networks, closed/known audience) for some instances (e.g., wedding) may also likely inhibit disclosure for some other instances (e.g., family conflict). This study therefore emphasizes a need to understand the interplay between audience and norms of a life event reportage in the online context. This can be studied via the lens of the socio-technical gap [5] to understand the fundamental discrepancy in facilitation of socio-technical systems — what individuals disclose online and what they disclose offline, and how the technical design of the systems may encourage one set of practices or goals over the other [5].

Contextual and Affordance Differences The findings show a contrast between social media disclosures and survey self-reports, which elicits a discussion of the respective modalities' affordances and context of use. We note that social media is naturalistic and largely recorded in-the-present unlike the survey which was retrospective and researcher-prompted; social media posting is also largely based on intrinsic motivation, whereas survey responses are driven by extrinsic motivation (e.g., monetary incentive). That said, both require individuals' active effort to be recorded. Accordingly, this study derives an interesting relationship with valence, significance and recency, and the ongoing nature of the life events — event attributes along which the reportage significantly differed (Table 6.13).

As discussed above, audience and impression management norms may make social media platforms to be less predisposed to sharing negative life events. However, why did the participants feel comfortable sharing negative life events with an audience of researchers? Compared to the social media audience that likely consists of strong and weak ties spanning online and offline interactions, researchers were strangers to the participants and comprised a smaller and likely perceived to be a more private audience than social media audience. These factors may have facilitated self-reporting of negative life events, free from concerns of stigma, social acceptance, or negative self-image.

Second, the findings support prior work that self-reported survey responses to likely be skewed to significant and recent events — significance and recency may cause disparities in emotional content, or salience, as these factors can change over time, especially after long time frames; emotional arousal may decay over time [148]. Extant literature lacks similar knowledge about online life events disclosures. This study contributes to this knowledge that significance and recency negatively associate with social media disclosures. The immediacy of active attention needed for a significant event may explain the lower likelihood of online posting. For instance, during a health emergency, an individual may not actively record a social media post, as the situation may demand attention to other more immediate, important needs. Again, in specific circumstances, significance of an event could be hard to understand in-the-present but may be realized only after a period of time [180], e.g., a dinner outing with a friend that becomes memorable after the friend's sudden, unexpected demise. Evolving significance can also lead to a different impression in memory, such as a case in this study when an individual posted about a vacation (with their significant other) on social media, but only self-reported about a breakup in the survey. Presumably, when the vacation began and was shared on social media, it initiated positive feelings, but after it ended with a breakup, the negative event stuck in the individual's memory.

Third, ongoing events are more likely to be shared on social media versus a survey, and that might relate to the social affordances of social media such as private messaging or an ability to write on someone's timeline; e.g., an individual in the process of moving between two places may feel like they can gather help, support, and advice relating to the move, as the event unfolds in real-time. These social affordances were absent in the survey conducted in this study, since the audience constituted the researchers, and the participants were recounting about life events from the past. Therefore, this study supports and provides complementary evidence to Haimson et al.'s work [285].

Ultimately, both in-the-present and retrospective perception of an event may depend on an individual's coping process [181, 657]. While validated surveys can measure how an individual coped with a traumatic or stressful life event, social media data can provide a stream of in-the-present recordings, e.g., this dataset contained a series of posts explaining the logistics, stress and support related to hospitalization process of an individual's child (identified as a continuous category). Surveys may also cause priming effects [555] — if a participant is inquired about a stressful life experience, they may undergo a psychological stress by re-thinking about those experiences. Considering these differences, this work shows that additional factors relating to events and individuals are important drivers of disclosures (and non-disclosures). To this end, this study also extends prior investigations that have examined the factors behind disclosure and non-disclosure on social media alone [18], by asking questions around how individuals arrive at decisions regarding which life event to disclose on social media versus self-report on a survey, and how these decisions straddle the contexts of use and affordances of the two modalities.

Design Implications

Considerable HCI research has sought to design, develop, and adapt platforms around life events like childbirth [164], gender transition [283, 284], and pregnancy loss [19]. Going beyond instances of specific life events, this study reveals that people not only share varied life events on social media, but also engage in selective sharing of life events, controlling for individual differences and event attributes. This study reveals a need to design for individuals and situations for both when disclosures do happen and when disclosures are withheld. Doing so necessitates closing the socio-technical gap per Ackerman [5].

Designing for Disclosure This study makes two design recommendations, first to scaffold the disclosure process itself, and second to make platforms and their algorithms sensitive to disclosures once they happen. Prior work reveals therapeutic and positive benefits of disclosure and expressive writing [51, 206], including benefits like finding an outlet for emotional release, self-acceptance, and solidarity with peers with similar experiences. This work finds that despite the occurrences of negative life events, individuals may not always disclose these events on social media, perhaps because of concerns noted as noted in the theoretical implications. As also noted by Andalibi et al. [19] and Ernala et al. [206], future research can therefore explore designing social media affordances that provide safe spaces for opening up for individuals with varied needs. This can include enabling individuals to create "trusted friend circles" based on various life event disclosures, e.g., a person may not be comfortable about sharing a work-related event but may be comfortable doing so with a set of trusted group of friends, therefore allowing targeted and staged disclosures [286, 644]. The findings suggest that users might be inhibited about disclosing negative or sensitive events. Users chose to not disclose certain events, despite Facebook providing audience control by design. To ease the process of recording an event privately or selectively, features may be included whose design and user experience are explicitly tailored to support the specific activity of recording life events, such as empowering users to define audiences and to limit the responses types about their life event, letting them take conversations to a different medium or outside of the platform, or having the provision of an expiration date on how long a life event may remain shared.

In addition, social media has shown promise as an intervention medium for crisis and wellbeing [124, 544]; we need to re-think alternative strategies for self-disclosures. For instance, to support individuals concerned about the public-facing nature of online platforms, tools may be built that emulate the benefits of personal blogging and journaling [131], to serve as a timestamped archive of one's thoughts and feelings around life events, empowering individuals to self-reflect traces of life. This can be a part of identity work or a part of memory work. This study also found that disclosure behavior may reveal an individual's momentary and longitudinal behavior, such as some disclosures being associated with momentary affective states (e.g., grief and joy), and others with lasting changes (e.g., moving to a different place). Consequently, this study suggests designing tools to provide supportive interventions around disclosures, including suggestions to rekindle interactions with social ties or recommending support communities.

On personal journaling, Facebook currently allows users to post and limit visibility to private. Some users send messages to themselves to record various events. However, none of these are by-design journaling interfaces. A recommendation could be an explicit private timeline space, where users can write private notes. Drawing motivation from smart journaling [201], such design can enable users to record life events, choose what to keep public and private, and also to toggle a private life event as public later in time. Further, platforms can consider designing with flexible anonymity, which can help break stereotyping or social expectations about social media posting of specific life events by particular demographics such as males and younger adults (as also seen in this study).

Next, as noted before, algorithmic content recommendation on social media is largely content and interests driven, showing personalized content based on individuals' interests and interactions with social ties. A lack of alignment of these recommendations with happenings in one's life, whether disclosed or undisclosed, can however have deep negative repercussions. This study noted an anecdote when algorithmic curation of Facebook feed was "inadvertently cruel" because it were not sensitive to an individual's life event [415]. Therefore, like prior HCI work [20, 91], tailoring recommendations to be inclusive and attuned to disclosed life events can increase the value people derive from these platforms. Literature notes that positive content can potentially benefit individuals to feel better in positive times, whereas supportive content may enable to feel comforted during adverse times [20, 447]. Such uses of social media can be promoted by designing life event-inclusive and -aware recommendation algorithms and affordances.

Designing for Non-Disclosure This study reveals that a “one size fits all” approach to scaffold online life event disclosures may not work. It matters not only that certain individuals choose not to disclose, but also that each event is associated with unique characteristics and circumstances. In particular, although this study did not solicit feedback from participants about why they chose to disclose or not disclose a particular event, certain demographic groups, such as males, older individuals, those low on agreeableness and extraversion personality traits were less inclined to disclosing online. Essentially, from a therapeutic perspective, the perceived efficacy of social media platforms as online social spaces to disclose life events, may vary across individuals. Despite having a Facebook account and using Facebook for other purposes, individuals may resist or reject using the platform to share personal happenings, as an individual choice, social practice, or the event’s temporality — a case for many of the participants. Scholars exploring technology non-use have found that disenchantment often stems from the perceived banality and inauthenticity of social interactions on social media platforms, particularly in contrast to offline communication [49]. Furthermore, some might feel socially disenfranchised to participate on a platform due to socio-institutional pressures, harassment, or social anxiety [459]. Because a disclosure might compromise an individual’s social network’s contextual integrity and the privacy expectations of other stakeholders of the life event [18], some of these factors behind non-use might play in this study as well. And yet, there were individuals who felt comfortable to self-report a life event on the survey, to a different social audience (of researchers), albeit smaller — indicating an implicit effort to weigh in the benefits and risks of disclosing life events on one modality versus another.

So how do we then design to accommodate the needs of these individuals with varying underlying decision-making processes around life event disclosures, and what would constitute an efficacious social media platform design for them? Given that this study reveals specific demographic differences among those who disclose and do not, how can design ensure that the groups who do not disclose are not marginalized?

Instead of designing only to encourage life event sharing on social media and risking "problematizing" the non-disclosers, we provide design suggestions drawing from scholars who have called for the role and perspective of the non-user to be recognized and valued [673]. First, platform designers need to account for social media non-use as a signal to modulate content recommendations. Essentially, design features may be built that allow individuals to curate or select what they would like to see and not see on the platform, depending on whatever their undisclosed current life event(s) might be. Second, drawing upon research on designing for technology non-use [49, 119, 481], platforms can accommodate alternative forms of participation for an individual, as a coping mechanism following an undisclosed life event, that does not involve being forced to deactivate or delete their social media account, or to stop social sharing and interaction altogether. For instance, individuals can switch platform settings to "no recommended content" and only visit parts of the site which they may feel are conducive to their current life circumstances. Broadly speaking, we draw from Baumer et al. [49], who noted that resistance to early telephone and electrical technology, particularly among rural populations, led producers to develop new designs and infrastructures better suited to rural life [348]. Similarly, this study encourages researchers and designers to make social media platforms life event-sensitive in a way that not only considers potential barriers preventing disclosures, but also provides agency in the decision-making processes behind non-disclosures.

Ethical Implications

Despite the best of intentions of a platform and designers to provide personalized content, works such as this, can lead to expectation mismatches, and individuals may perceive intrusiveness and dissatisfaction about such algorithmic content curation without consent [214]. Further, identifying life event disclosures on social media can lead to other ethically questionable consequences such as targeted advertising [319], including compromised privacy, defying expectations, and damaging relationships — reminiscent of the case of the woman whose pregnancy was discovered by a supermarket chain without her knowledge [301]. Although personalizing ads around positive events (e.g., new home, wedding) may bear both business and individual advantages, the same around negative life events can not only exacerbate an individual's situation and wellbeing, but also can be deemed unethical and intrusive [375].

Furthermore, people's online disclosures of life events can be (mis)used to infer high-risk decision outcomes in one's offline life such as job, insurance coverage, financial support, or obtaining a property mortgage. At the other end of the spectrum, when people do not disclose their life events, it might prevent such misuse, but they may be disadvantaged in deriving the benefits that disclosing individuals might be able to derive from the platform, such as access to support, social capital, or social approval. From a social computing standpoint, both disclosures and non-disclosures of life events on social media can lead to forming new social conventions and norms on the platform with repercussions on an individual's life, e.g., research already notes the positivity bias on social media [90], and non-disclosure of negative events may make people feel worse when they experience a negative life event. Overall, these ethical complexities call for better understanding and guidelines regarding what platforms owners and decision makers can and should do with people's (non)-disclosures of life events, for what purpose, and the extent to which transparency is baked into these uses.

CHAPTER 7

OBSERVER EFFECT IN SOCIAL MEDIA BEHAVIOR

The previous chapters reveal the potential of social media in inferring psychosocial dynamics of individuals and collectives in situated contexts [155, 528, 530, 531, 536, 544]. We also see how we can overcome the limitations of social media data by using complementary multimodal sensing data, and how it is important to introspect and interpret these data-driven assessments given the real-world implications. We note that most of the studies on social media based assessments rely on *retrospectively collected* data. The extant literature informs and motivates human-centric technologies to improve wellbeing in real-time and prospective settings. However, prospective use of social media based assessments would bring in new challenges, and which are yet to be studied and addressed.

In contrast to retrospective settings, prospective social media studies can likely suffer from "observer effect", or the phenomenon that individuals might deviate from their normative behaviors, attributed to the awareness that they are being watched or studied [458, 609]. The social ecological model theory posits that human behavior is embedded in the complex interplay between an individual, their relationships, their communities, and societal factors [102]. While this theory explains the potential of social media as a viable sensor of human behavior, it also points out a caveat — the observers, who are also a part of a subject's ecology, may affect the subject's behavior (or in other words, the observer effect). This chapter explains my investigation towards measuring and studying observer effect in social media sensing studies on wellbeing.

Therefore, if we envision a future with social media technologies for wellbeing interventions, we should ask how these algorithms would perform if we conduct *prospective* data collection? The algorithms that are built on retrospectively collected data may not be as effective in prospective settings. Despite the potential and the evidence, the validity of adopting computational social science on big data has been critiqued regarding its applicability in real-world and in practice [73, 366, 525, 624]. Ruths and Pfeffer (2014) pointed out that these studies may misrepresent or be ineffective in the real-world [525]. Lazer et al. (2014) noted how the Google Flu predictor algorithm that uses google search behavior, over-estimated the number of flu visits in real-time, even though it performed exceptionally well on historical data [366].

The validity and in-practice reliability of human-centered big data technologies suffer due to the unpredictability and complexity in human behavior along-with unaccounted confounds [452]. Again, the ecological validity of these technologies remain unattested because observer effect is not typically accounted for. As posited in psychology and social science literature, "observer effect" is the phenomenon that people are likely to alter their behavior with the awareness of being monitored or observed [609].

Although social media is a passive sensor in terms of data collection, it is dependent on the active creation by individuals in their self-initiated desire to engage on a social media platform. Given that social media behavior is a form of intentional and conscious behavior, or a behavior that individuals can alter at their will in case they feel "observed", *observer effect remains a basic unexplored phenomenon that may bias observations from social media sensing of human behavior.*

Historically, observer effect has been hard to study because researchers could only access data generated after participant recruitment. However, measuring observer effect necessitates access to a subject's normative and non-observed behavior, or the counter-factual how the subject would have behaved without the presence of an observer. In contrast, social media is a longitudinal and historical data stream, and allows access to years of an individual's behavior on the platform, including pre-enrollment data. This provides the opportunity to build behavioral models on normative or "non-observed" behaviors. By definition, observer effect can be avoided if studies are conducted without the awareness of the participants. For instance, retrospective and observational data is one such solution (as also described in the previous chapters). However, these settings are not always applicable, such as in identifying historically unobserved behaviors or conducting proactive interventions. Alternatively, researchers can conduct experimental studies without participants' awareness. For example, the Facebook emotion contagion study did not seek the consent of participation and did not make the users aware that their social media feeds were modified for experimental purposes [357]. Although this study successfully uncovered valuable insights about people's affective behaviors on social media, it was heavily critiqued on ethical grounds [331]. Again, only a select few researchers have privileged access to conduct such studies, and those who do, may be institutionally constrained in studying or revealing certain phenomena [73, 394, 624].

As boyd and Crawford (2012) noted, these practices (of conducting studies and experiments without participant awareness) can reinforce the troubling perception of these technologies as "Big Brother, enabling invasions of privacy, decreased civil freedoms, and increased state and corporate control" [73]. Moreover, the best practices of research entail obtaining informed consent from participants, making them aware of what data they share and how it will be used. But this also increases the likelihood of observer effect as participants may self-alter their behaviors as postulated in theories such as social desirability, psychological reactance, self-presentation, self-monitoring, and reasoned action, to name a few [249, 390, 587, 598]. A better understanding of observer effect in social media sensing studies would make researchers aware of what they can expect and make us think towards designing and developing measures to correct and account for this effect in future study designs and observations [525].

Motivated by the above factors, I examine observer effect on social media behavior, by asking the following research questions:

- **RQ1.** What is the prevalence and degree of observer effect in social media behavior?
- **RQ2.** How psychological traits of individuals explain their likelihood to show observer effect in social media behavior?

My examination is situated within a year-long multimodal sensing study, where participants consented social media data, particularly Facebook. This work draws motivation from person-centric studies in psychology research of clustering individuals on similar psychological traits, and from causal-inference research in minimizing the confounds. I adopt time-series and statistical modeling to measure how individuals deviated from expected behaviors after enrolling in the multimodal sensing study, or their awareness of being "sensed". I operationalize observer effect in two dimensions of social media activity, 1) behavioral changes, and 2) linguistic changes. Within behavioral changes, I measure the deviation in quantity and verbosity of posting behavior, and within linguistic changes, I measure the topical and psycholinguistic changes. This study reveals that observer effect indeed occurs, but its occurrence varies across participants in various ways. For instance, individuals with high cognitive ability and low neuroticism show an immediate decrease in social media posting after enrollment, but their behaviors get closer to normalcy over time. In contrast, individuals with high openness do not show any immediate changes in posting quantity, but their posting significantly increases over time. Linguistically, most individuals show a decreased use of first person pronouns, which reflects reduced sharing of intimate and self-attentional content. While some individuals increase posting about public-facing events, others increase posting about social and family gatherings. Finally, I explain the behavioral changes with respect to psychological traits in a theory-driven fashion.

Theoretically, this work advances our knowledge about how individuals varying on psychological traits could differently change social media behaviors. These behavioral changes are explained by social science and psychology theories, including self-monitoring [587], public self-consciousness [26], and psychological reactance [78]. Methodologically, this work contributes a computational and causal framework of modeling and inferring observer effect in human-centered studies in general, and social media sensing in particular. This work provides insights regarding whether observer effect occurs, how long does it last, and how does its occurrences vary across participants. I discuss the implications of this work in recommending strategies to correct biases due to observer effect in social media sensing studies.

7.1 Background and Related Work

7.1.1 Observer Effect in Research

Definition

Literature in psychology and social science posits that people are known to alter their behavior with the awareness of being monitored or observed. This phenomenon is known as the "observer effect" or the "research participation effect" or the "experimenter effect" [512]. This effect is also popularly called the "Hawthorne effect" as historically this effect was first seen in a series of experiments between 1924 and 1939 in the Hawthorne Works plants [118, 513]. Observer effect has been commonly cited to affect the reliability of observations in studies [343]. For our purposes, we use the following definition as outlined in a systematic review done by McCambridge et al. (2014).

> "The Hawthorne effect concerns research participation, the consequent awareness of being studied, and possible impact on behavior." — McCambridge et al. 2014 [609]

Given that there are several arguments around the use and appropriateness of the term and context corresponding to "Hawthorne Effect" [118, 121, 450, 609], our work adopts the term, "observer effect" for disambiguity and consistency purposes.

Hawthorne Studies

The Hawthorne studies were originally conducted at the Hawthorne Works Plant near Chicago where six studies were conducted between 1924 and 1933. These studies were longitudinal in nature spanning between several months and several years. These studies were designed to understand and improve worker productivity at the plant. Initially, these studies were investigating the effects of different levels of illumination on productivity.

However, these experiments inconclusively found that no matter what timing or kind of illumination was used, worker productivity always increased. This shifted the interest of researchers that these changes in behavior (or productivity) was possibly not due to the changes in illumination level per se, but due to the awareness of the workers of being monitored or “observed”. This shift of interest led to examining a variety of social and psychological factors that contribute to changes in human behavior [121].

A number of social science and psychology theories have been proposed in the last century that examines human behavior, and explains behavioral change with respect to observer effect in different settings. Guerin (2010) reviewed research on behavioral change in the presence of others (social presence), and postulates the phrases of “social facilitation” and “social inhibition” as opposite effects on a worker’s performance. This effect is also described as a form of “reactivity” as individuals modify an aspect of their behavior in *response* to a phenomenon (awareness of being observed) [342, 598]. Based on the social desirability theory, conformity and social desirability considerations can lead behavior to change in line with these expectations [121, 265]. Observer effect is also frequently studied in epidemiological and clinical studies in order to minimize confounds in findings [118, 225]. Observer effect has also been attributed to affect methodologies such as field observations and ethnography [367], and is considered to be one of the biggest challenge and long described as the “Achilles heel” of participant research [458, 584].

Researchers have been interested in limiting experimenter-observer interactions that may cause observer effect [512]. Longitudinal studies have shown promises of mitigating such effects because participants either adapt to normalcy or become less aware of being observed over time [678]. Alternatively, observer effect can be considered to be a strength (instead of limitation) in certain settings, as participants may behave more ethically, conscientiously, and efficiently [427]. Observer effect is not necessarily viewed as surveillance can also contribute towards increased inter-accountability due to co-presence [52].

However, none of the above may apply in our particular setting of social media sensing.

Social media data, by its very nature and strength, is created by an individual in naturalistic settings by their self-motivated and self-initiated will, and is collected passively and unobtrusively. An individual who consents to sharing their social media data, may not actively feel aware of being observed all the time. This awareness might influence certain behavioral amendments that essentially normalize over time, or a process known as *habituation* in behavioral sciences [117]. To understand the likelihood and extent of observer effect on social media behavior, we examine social media behavior following enrollment in a year-long multisensor study. By adopting a causal-inference approach that measures the deviation in an individual's post-enrollment social media behavior from their expected behavior, we minimize the confounds and delineate observer effect from other behavioral aberrations at chance. This work makes theoretical contributions to the general interest of of understanding and leveraging social media sensing for human behavior.

7.1.2 Behavior Change on Social Media

Over the years, people's social media behavior has been studied in a variety of ways, spanning across prediction and inference studies on information dissemination, political interests, stock market, and wellbeing [2, 71, 106]. The growing evidence of the relationship between human behavior, psychology, and language allows us to infer these behavioral changes when we analyze longitudinal social media data. Similar to our physical world, people's online presentation is a factor of their social network [189, 305]. Guillory and Hancock found that the public-facing nature of platforms such as LinkedIn influences an individual's accountability and reduces deception in their self-description of their professional portfolio, which also aligns with Donath's early research on identity and deception in online spaces [183]. Reinecke and Trepte (2014) found that social media provides an environment for online authenticity, an in fact authentic self-presentation contributes towards positive psychological wellbeing [501]. Similarly, a body of literature revealed evidence regarding how social media facilitates candid self-disclosure for an individual [165, 506].

In the specific interest of human-centered studies of wellbeing dynamics, prior work has studied behavioral changes on social media in several contexts. For example, De Choudhury et al. (2013) examined social media behavior changes around a major life event, particularly postpartum changes in behavior and mood of new mothers along the dimensions of social engagement, emotion, social network, and linguistic style [162]. Golder and Macy (2011) studied the variability in mood and sentiment in weekends and weekdays. Other longitudinal studies have examined behavioral changes around exogenous or endogenous, anticipated or unanticipated events, e.g., antidepressant use [540], counseling advisories [544], alcohol and substance use [346, 380], diagnosis with mental health conditions [206, 277, 278], suicidal ideation [167], and so on. Relatedly, Ernala et al. (2018) adopted the Social Penetration Theory to operationalize intimacy of self-disclosure and studied how it varies with respect to audience engagement on social media [205].

Researchers have also explored behavioral changes around topics related to observer effect, such as privacy. Back in 2014, when Zhang et al. studied "creepiness" and privacy concerns related to social media use, they found concerns shifting from boundary regulation to behavior tracking by social media platforms for targeted advertising [682]. However, social media- and the web- based behavioral inferences have evolved since then, and have also come under ethical and political scrutiny for privacy breaches such as the Cambridge Analytica scandal on Facebook [94]. This has also renewed attention to the challenges that may arise when data is appropriated for surveillance by stakeholders, e.g., workplace surveillance [236]. At the same time, concerns related to audience, boundary, and disclosure regulations are evident on social media, people want themselves to be viewed in particular ways across different audiences [189, 338, 403, 623]. As per Goffman's theory of self-presentation, individuals may present two kinds of information (including on social media) — one that they intend to "give off", and one that "leaks through" without any intention [249, 421, 666]. One strategy of boundary regulation, self-censorship is known to be prevalent on social media [151, 403]. Self-censorship occurs when social media users prevent themselves from posting or conducting a behavior despite a self-initiated initial desire to do so [151]. For example, Wang et al. studied self-censorship with respect to regretful thoughts [654]. Also, privacy-concerns may lead to changes in social media behavior, in terms of presentation, censorship, and information sharing [6, 623]. However, researchers have also found an apparent "privacy paradox", i.e., despite the awareness privacy concerns, individuals may share more personal information on social media [40]. This shows that people's social media behavior has been found to be complex, which also depends on each individual's personality, perceptions, and beliefs as well as external factors [273, 500].

In addition, the earlier chapters also revealed a variety of factor may describe why and how an individual self-describes themselves on a social media platform (e.g., LinkedIn and Facebook) [536, 538]. While social media data is a useful signal to analyze behavioral changes, people's perceptions about the use of social media may significantly affect their behavior. The current study explores this phenomenon by leveraging longitudinal social media data to delineate effects of observer effect on people's social media behavior. While no theoretical framework can be directly adopted to understand behavior changes around observer effect, in the next section I review suitable behavior change theories that help in explaining social media behavior change in the context of observer effect.

7.1.3 Theories of Behavior Change in Research

Social scientists and psychologists have proposed numerous theories related to behavioral change. The socio-cognitive theory adopts an agentic perspective to human development, adaption, and change by distinguishing three modes of agency, personal, proxy, and collective [38]. The situated identity theory states that relevant cues in behavioral settings are first translated to identity potentials which provide the basis for specific behavioral choices [13]. Self-consciousness is another construct that may influence one's strategic self-presentation [26]. Again, the concept of psychological reactance describes that individuals have certain freedoms regarding behaviors, which if reduced or threatened, they react in

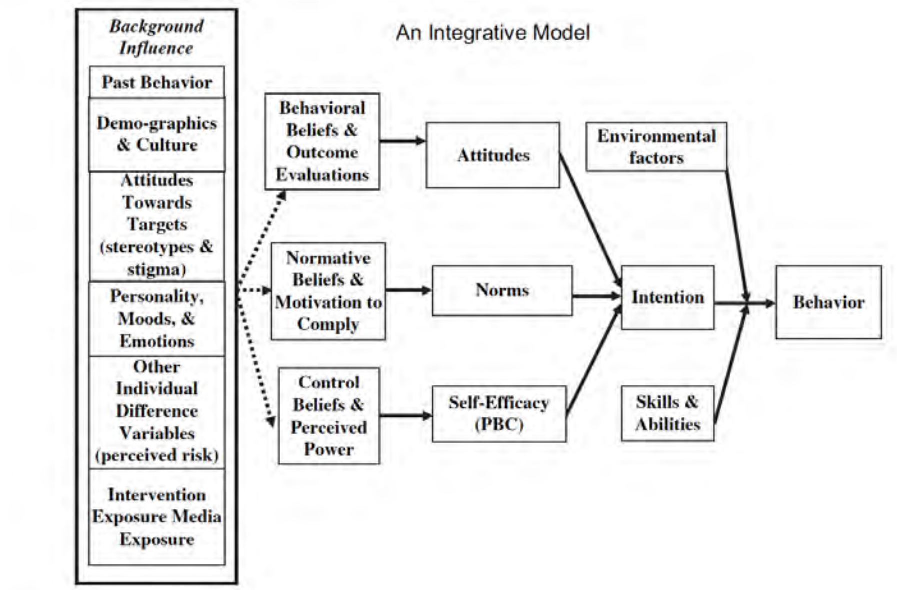


Figure 7.1: Integrative model on behavior change as proposed in [218]

order to regain them [78]. Introduced by Snyder, the concept of self-monitoring posits that people self-monitor their self-presentations, expressive behavior, and non-verbal affective displays [587]. Self-monitoring can be considered to be a form of personality traits that regulate behavior to accommodate social situations [587].

Further, Fishbein and Capella 2006 note, “Although there are many theories of behavioral prediction such as the Theory of Planned Behavior [8, 9], the Theory of Subjective Culture and Interpersonal Relations [621], the Transtheoretical Model of Behavior Change [490], the Information/Motivation/Behavioral-skills model [219], the Health Belief Model [53, 516, 517], Social Cognitive Theory [35, 37], and the Theory of Reasoned Action [217], a careful consideration of these theories suggests that there are only a limited number of variables that must be considered in predicting and understanding any given behavior [216]” [218]. The integrative model as per Fishbein (2000) attempts to bring together a number of theoretical perspectives and is presented in Figure 7.1 [218].

This dissertation adopts the above theories and variables to interpret and explain observer effect in social media behaviors. After quantifying the deviation in post-enrollment actual and expected behaviors, I investigate how people’s individual differences could likely explain the behavioral changes, by situating in the above theories. Further, the insights from this work leads to generating hypotheses which can be tested and evaluated in future research.

7.2 Study and Data

Like the studies in the previous two chapters, the data for studying observer effect comes from the Tesserae project (section 4.1). As described before, the participants responded to initial survey questionnaires related to demographics, and trait-based measures relating to personality, affect, sleep, and executive functions. The participants were requested to remain in the study for either upto a year or through April 2019. The study enrollment was conducted from January 2018 through July 2018.

7.2.1 Social Media Data

The Tesserae project asked consented participants to authorize their Facebook data, *unless they opted out, or did not already use Facebook*. The enrollment briefing and consent process explicitly explained that the study participation did not necessitate them to use social media in a particular fashion, and they were expected to continue with their typical social media use. The participants authorized access to social media data through an Open Authentication (OAuth) based data collection infrastructure developed in Saha et al. [528].

Given that Facebook is the most popular social media platform [262] and its longitudinal nature has enabled several of human behavior [166, 664], it suits our problem setting of understanding observer effect in social media behavior. Out of the total 572 participants who provided access to Facebook data, 532 made at least one post on their Facebook timeline. Table 7.1 summarizes the Facebook dataset of Tesserae participants, and I find that there is roughly 82 months data per participant in the pre-enrollment period, and roughly 5 months data per participant in the post-enrollment period. Among these participants, we apply a filter of participants with at least 60 days of post-enrollment data, leading to 316 participants, whom we examine for observer effect in the rest of the study.

Table 7.1: Summary of pre- and post- enrollment Facebook datasets.

	Before Enrollment		After Enrollment	
Type	**Range**	**Mean**	**Range**	**Mean**
Posts	26-4,472	865	8-964	101
Comments	34-10,228	1,593	5-1,104	175
Likes	62-52,139	6,536	15-4,540	940
Duration (months)	0-160.27	82.52	0-12.87	4.59

Table 7.2: Summary of demographics and individual differences of 316 participants whose data is being studied for observer effect.

Covariates	Value Type	Values / Distribution
Demographic Characteristics		
Gender	Categorical	Male | Female
Age	Continuous	Range (21:63), Mean = 36.36, Std. = 10.28
Education Level	Ordinal	5 values [HS., College, Grad., Master's, Doctoral]
Job-Related Characteristics		
Income	Ordinal	7 values [<$25K, $25-50K, ... , >150K]
Tenure	Ordinal	10 values [<1 Y, 1Y, 2Y, ... 8Y, >8Y]
Supervisory Role	Boolean	Non-Supervisor | Supervisor
Cognitive Ability (Shipley)		
Fluid (Abstraction)	Continuous	Range (5:24), Mean = 16.93, Std. = 2.95
Crystallized (Vocabulary)	Continuous	Range (18:40), Mean = 33.55, Std. = 3.57
Personality Trait (BFI)		
Openness	Continuous	Range (2.2:5.0), Mean = 3.82, Std. = 0.59
Conscientiousness	Continuous	Range (1.9:5.0), Mean = 3.91, Std. = 0.64
Extraversion	Continuous	Range (1.7:5.0), Mean = 3.42, Std. = 0.66
Agreeableness	Continuous	Range (2.3:5.0), Mean = 3.93, Std. = 0.55
Neuroticism	Continuous	Range (1.0:4.6), Mean = 2.46, Std. = 0.78
Affect and Wellbeing		
Pos. Affect	Continuous	Range (13.0:49.0), Mean = 34.15, Std. = 5.82
Neg. Affect	Continuous	Range (10.0:40.0), Mean = 17.06, Std. = 4.88
Anxiety	Continuous	Range (20.0:67.0), Mean = 37.05, Std. = 9.28
Sleep Quality	Continuous	Range (1.0:16.0), Mean = 6.72, Std. = 2.51

7.2.2 Self-Reported Survey Data

Tesserae project's enrollment process included initial demographics surveys (age, gender, education, income, etc.), and surveys of self-reported psychological constructs as explained in section 4.1

7.2.3 Preliminary Analyses

First, I conduct some feasibility and preliminary tests on our data for our study.

Statistical Power

Power analysis in statistics estimates the minimum sample size for a study to make significant inferences on a given population [607]. Likewise, I use power analysis to examine if this study has sufficient sample size of participants to make reasonable inferences about the population. This study's participant pool belongs to information workers in the United States. According to U.S. Census Bureau, a rough estimate on the number of information workers in the U.S. is 4.6 million [503]. I calculate a sample size that is representative of this population with a 95% confidence interval and 5% margin of error, this comes out to be a sample size of 385. Given that the net social media sample size is 574 participants, out of which, usable data for studying observer effect is for 316 participants, this study assumes to have a reasonable sample of information workforce in the United States.

Quantity of Posting

Posting behavior is a prominent social media behavior that has revealed significant signals of human behavior in prior work [162, 206, 539, 544]. I measure the average posting behavior of the participants over time and around their enrollment in the study. Figure 7.2 shows the daily average posting behavior of the participants relative to the day of enrollment, where day=0 corresponds to the enrollment day for the participants. We notice an apparent bump in the average number of posts per day post-enrollment in the study.

Expressive Behavior

I also examine the changes in expressive behavior of the participants. For this, I use the psycholinguistic lexicon LIWC [613] to obtain the psycholinguistic changes in the participants' post following enrollment in the study. Figure 7.3 reports the effect sizes comparing pre- and post- enrollment normalized use of psycholinguistic categories across the participants. A positive effect size indicates greater use of the category post-enrollment, whereas a negative effect size indicates lower use in the post-enrollment period. Effect size (Cohen's d)

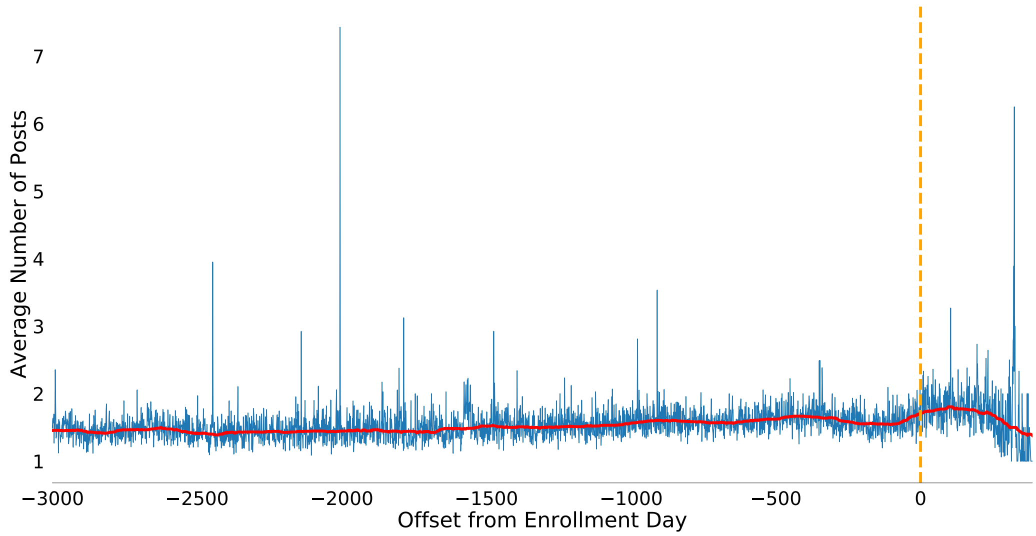


Figure 7.2: Average number of posts per day across all participants on relative offset from their day of enrollment. Day 0 indicates the day of enrollment.

is considered to be a significant difference if its magnitude is greater than 0.15. We find that at an aggregated level, multiple psycholinguistic categories show significant changes. For example, considering pronoun use, first person pronoun use decreases, which might indicate a decreased sharing of intimate content and decreased self-attentional focus [131]. In contrast, the use of first person plural, second, and third person pronouns increase. We also find a decrease in the use of cognition related words (such as cognitive mechanics, discrepancies, inhibition, negation, etc.). We also find a significant decrease in affective categories of anger, sadness, and swear.

The above preliminary analyses indicate certain changes people's behavioral and expressive social media use following enrollment in the study at an aggregated level. This motivates us to examine the changes in a much more rigorous and robust fashion. Given that not all individuals are the same, this study borrows from person-centric approaches to examine the changes in cohorts (clusters) of similar individuals on psychological constructs [154].

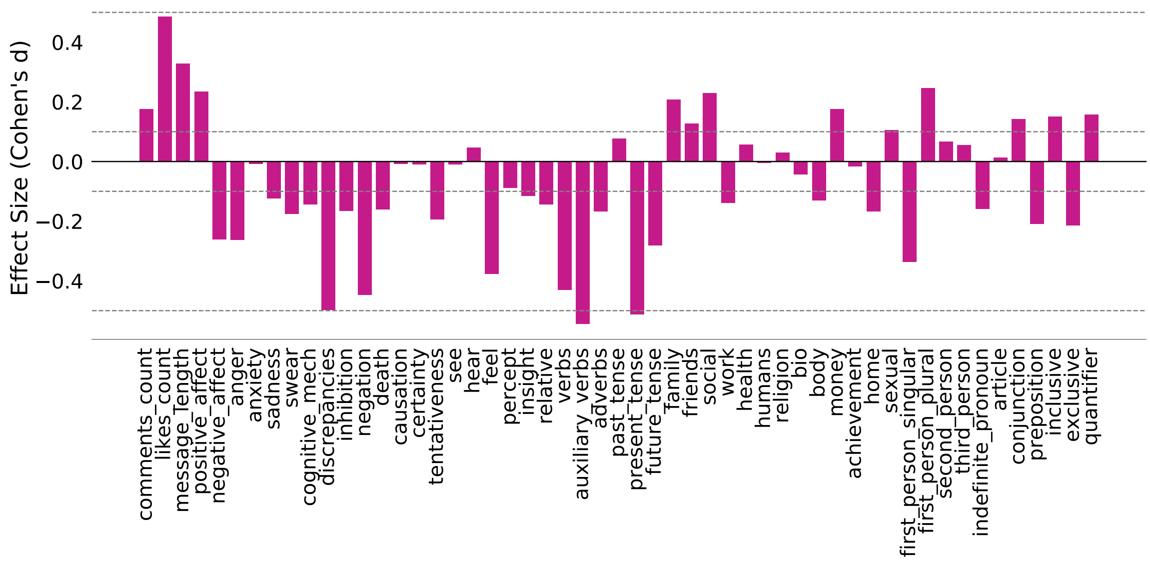


Figure 7.3: Effect size (cohen's d) comparing before and after enrollment datasets of users across psycholinguistic attributes. A positive cohen's d indicates that post- enrollment data Cohen's d magnitude smaller than 0.20 is considered to be small difference.

7.3 Methods

This study operationalizes observer effect as the deviation in actual post-enrollment social media behavior from expected (or normative) behaviors. In particular, I measure the changes in two dimensions of — 1) behavioral changes (posts made and engagement sought), and 2) linguistic changes (topics and psycholinguistics). I examine these changes in a person-centric approach of clustering individuals on psychological traits. This section describes my approach of clustering individuals followed by measuring the dimensions of social media behavioral change.

7.3.1 Clustering Participants on Intrinsic Traits

Typically, prediction models are studied on the entire dataset of participants, also termed as variable-centric or generalized prediction approaches, where a single model is built for the entire training data available. However, in contrast to many other datasets, social media data presents unique challenges, as it is sensitive to people's social media use and

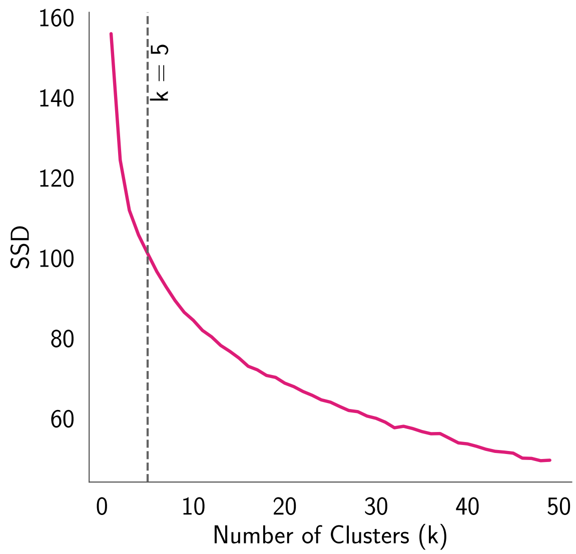


Figure 7.4: Elbow plot to estimate the optimal number of clusters by varying number of clusters (k) and mean sum of squared distances to the cluster centroids (SSE).

may significantly vary across individuals. Although personalized approaches could help overcome this challenge [391, 527], it is hard to conduct personalized examinations on social media data, because this data suffers from sparsity issues, compromising the statistical power. Therefore, drawing on prior work [154, 532], I adopt a middle ground that balances the trade-offs between too personalized and too generalized models. In particular, I cluster individuals with similar traits, and then examine the outcomes per-cluster. This approach is known to balance both between-individual homogeneity and within-individual heterogeneity in our behaviors [532].

I adopt a k-means clustering approach to cluster individuals on intrinsic traits as collected via ground-truth surveys (personality traits, cognitive ability, affect, anxiety, and wellbeing). I employ the elbow-heuristic to obtain the optimal number of clusters (k) in our approach [552]. Figure 7.4 shows the elbow plot of mean sum of squared distances to the cluster centroids with respect to the number of clusters (k), roughly estimating an optimal number of clusters at k=5.

Consequently, I conduct k-means (k=5) clustering on the intrinsic traits of individuals to cluster the initial 513 individuals in the dataset. We obtained five clusters ($\mathtt{C}_0$ to $\mathtt{C}_4$),

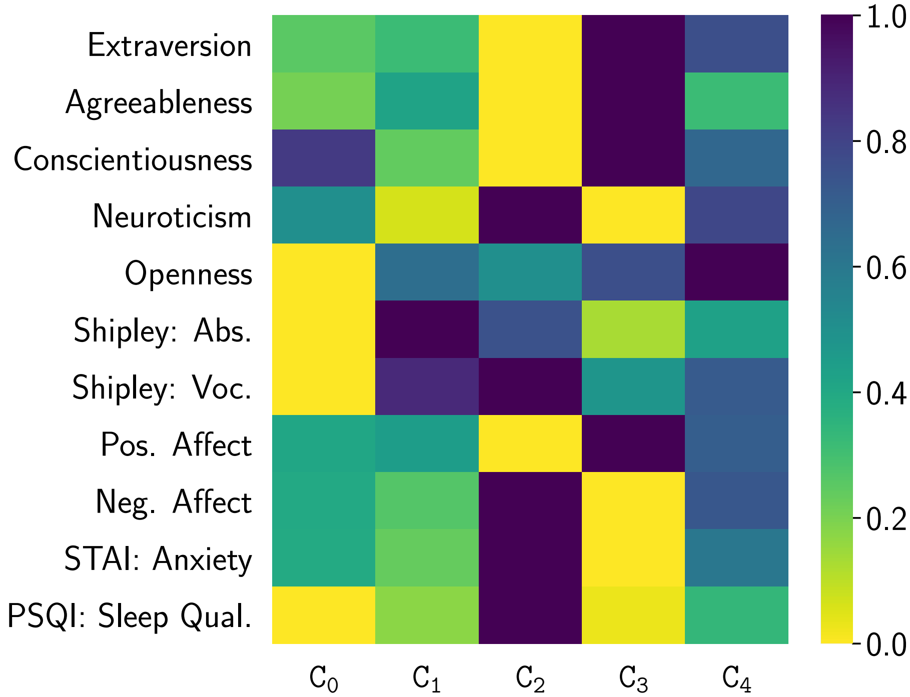


Figure 7.5: Distribution of traits across clusters of individuals.

Table 7.3: Summary of descriptions of clusters on psychological traits.

Cl.	N	Characteristics
C_0	60	High (Conscientiousness, Sleep Quality), Low (Openness, Cognitive Ability)
C_1	66	High (Cognitive Ability), Low (Neuroticism)
C_2	44	Low (Extraversion, Agreeableness, Conscientiousness, PA, Sleep Quality), High (Neuroticism, Cognitive Ability, NA, Anxiety)
C_3	97	High (Extraversion, Agreeableness, Conscientiousness, PA, Sleep Quality), Low (Neuroticism, NA, Anxiety)
C_4	49	High Openness

containing 93, 115, 70, 152, and 83 members respectively. Figure 7.5 shows the average distribution of the traits and Table 7.3 summarizes the characteristics of the five clusters.

7.3.2 Measuring Behavioral Changes

Measures to Quantify Behavioral Changes

I examine the participants' post-enrollment behavioral changes on social media. This includes the changes in quantity and verbosity of the posts. Additionally, social media behavior is also characterized by social networking and engagement received from others. Therefore, I also examine the changes in the quantity of likes and comments received on the participants' posts. I explain the measures below.

Posting Behavior. We examined social media posting behavior in two measures — 1) *Quantity of posting*, i.e., the daily average number of posts, and 2) *Verbosity of posting*, i.e., the daily average number of words.

Engagement Received. I examine the engagements received on social media posts, in terms of 1) *Likes*, i.e., the daily average number of likes received, and 2) *Comments*, i.e, the daily average number of comments received.

Modeling and Quantifying Behavioral Changes

Drawing on interrupted time series and synthetic control based causal approaches [48, 408], I compute the deviation in actual behavior from the expected behavior of the participants as modeled on their historical behavior. For each cluster, I build autoregressive models (ARIMA) to predict post-enrollment expected behaviors of the participants. I train the models on the pre-enrollment data, using a $80:20$ split (80% for training and 20% held-out for testing), and applied grid search to optimize for the best parameters of the time series prediction models. The models are evaluated on the 20% held-out data, as symmetric mean absolute percentage error ($\mathtt{SMAPE}$) which quantifies errors in the range of 0 to 100, where lower values indicate a better predictive model. Besides, I measure the statistical significance in difference of actual and expected behavior using paired t-tests and effect size (Cohen's d).

Conducting Placebo Tests

Further, I need to ensure that the effects observed in the study are most likely an artifact of study enrollment, and not due to other confounds or at chance. For this purpose, I conduct placebo tests drawing on permutation test approaches from prior work [17, 537]. Within the pre-enrollment data, I permute (randomize) on several *placebo* dates. I assign 150 placebo dates, and repeat the above time series comparison around the placebo dates — for every placebo date, I compute the t-tests in the post- placebo date actual and predicted time series data. Then, over all the permutations of placebo dates, I compute the probability (p-value) of significant differences around placebo dates. A p-value lower than 0.05 would reject the null hypothesis that the significance is by chance, also revealing the credibility about any significant changes observed around the (real) enrollment date.

7.3.3 Measuring Topical Changes

Topics are a useful means to understand the content of people's social media expressions [111]. I conduct topic modeling in our dataset to examine how the prevalence and diversity of topics evolve following study enrollment. First, to automatically extract topics, I employ the widely adopted Latent Dirichlet Analysis (LDA) on the dataset [68]. LDA is known to produce stable and interpretable topics, and has often been used in social media and human behavior research [111, 206, 504].

Building Topic Models

To identify the optimal number of topics in our dataset, we draw on the recommendations from Wallach, Murray, Salakhutdinov, and Mimno and Chang, Boyd-Graber, Wang, Gerrish, and Blei. That is, I vary the number of topics upto 25, and semi-automatically evaluate the quality of topic models, by combining the use of topical coherence scores as well as manual evaluations. Topical coherence score quantifies the degree of semantic similarity between high scoring words within a topic [413]. Figure 7.6 plots the coherence scores on varying the number of topics from 2 to 26, suggesting that the highest coherence is achieved at around the number of topics (n) as 10. In addition, I, along with two other collaborators in the research team, manually evaluate the topical distribution for n=8, n=10, and n=12. We find the topical distributions at n=8 and n=10 to be less semantically coherent, with a substantial increase in noisy keywords. Therefore, as guided by both coherence scores and manual examination, I use topic modeling for n=10 topics for our ensuring analysis.

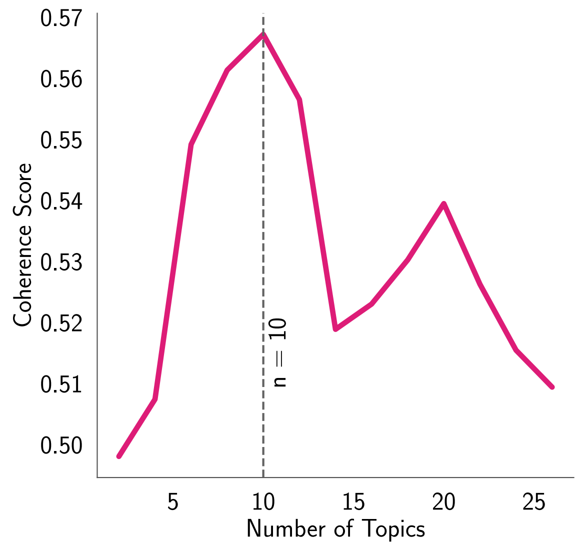


Figure 7.6: Topical coherence scores on LDA topic modeling with varying number of topics.

Interpreting Topics

After building the topic models, I assign interpretable labels to topics and keywords. For this purpose, three members of the research team (including me) design an interpretive annotation to identify coherent themes in the keywords per topics. The topics are first inductively and independently coded with implied themes. Then the codes are compared and agreed upon to assign final thematic labels per topic. The thematic category of a topic is implied from the within-topic coherence and between-topic separation of keywords. These themes are 1) *Travel and Locations*, 2) *Food and Drinks*, 3) *Holiday Plans*, 4) *News and Information*, 5) *Work-Life Balance*, 6) *Family Gathering*, 7) *Social and Sports*, 8) *Greetings and Celebration*, 9) *Friends and Family*, and 10) *Acivities and Interests*. Table 7.4 shows the 10 thematic categories and top occurring keywords per topic, along with example paraphrased post from our dataset.

7.3.4 Measuring Psycholinguistic Changes

Another important dimension to understand people's expressiveness and social media behavior is through psycholinguistics, which is known to associate with psychological

Table 7.4: Thematic categories of topics identified in our dataset.

Theme	**Topic Words**	**Example post**
Travel & Locations	country, green, baby, miss, right, chicago, sad, need, let, denver, mean, airport, hello, way, win, begin, yum, national, cubs, joanie	*Smiles all around after a good ATD conference together in Denver.*
Food & Drinks	lol, new, ready, room, sweet, boy, getting, waiting, finally, time, chicken, need, delicious, chicken, got, cheese, food, beer, gotta, yeah, guess	*Chicken on the grill, beef roast on the cutting board, regular and sweet potatoes in the oven. Guess who's not cooking tomorrow!*
Holiday Plans	christmas, school, vote, today, true, high, trip, look, season, awesome, johnson, merry, news, summer, party, check, raise, mom, family	*Morning hike, trip to the beach, and relaxing at our rental!*
News & Information	like, people, time, things, trump, think, watch, know, looks, right, got, thing, want, need, good, going, bad, stop, run, better, org	*Climate models want to change the way we live ... should we listen? It's a short video, watch it.*
Work-Life Balance	home, work, day, got, yes, new, today, time, tomorrow, little, house, going, like, car, snow, hours, bed, dog, night, way	*After work. Only one thing on my mind.*
Family Gathering	good, morning, great, night, fun, day, time, weekend, dinner, week, friday, today, tonight, party, work, family, team, going, view, date, girls, weekend	*Had a great visit with Otto & family!*
Social & Sports	game, want, tony, retweeted, play, south, come, bend, dame, notre, it's, tulio, tickets, world, need, free, shit, dace, wants	*Watched my team in India play a friendly cricket match last night and got a lesson on the difference between batting in baseball versus cricket.*
Greetings & Celebration	day, happy, love, birthday, wedding, today, anniversary, halloween, disney, beautiful, mom, http, best, little, year, life, wish, challenge, thank	*Wishing my beautiful daughter a wonderful birthday. Love you baby girl.*
Friends & Family	years, time, love, family, friends, year, life, thanks, kids, amazing, best, know, today, old, wait, great, ago, days, help, people	*Enjoying St Helena, brunch and wine tasting with my son and friends.*
Activities & Interests	like, read, years, wow, know, love, good, think, people, music, interesting, post, facebook, ago, copy, itś, wheels, place, favorite, book	*First book Iv́e read in a long time that I couldn't put down. The Life We Bury*

states and attributes [166, 564]. To conduct psycholinguistic analysis on our dataset, I use the widely adopted lexicon of Linguistic Inquiry and Word Count (LIWC). LIWC is a psychologically validated lexicon that allows to categorize the pre- and post- enrollment social media data into psycholinguistic categories of: 1) *affect* (categories: anger, anxiety, negative and positive affect, sadness, swear), 2) *cognition* (categories: causation, inhibition, cognitive mechanics, discrepancies, negation, tentativeness), 3) *perception* (categories: feel, hear, insight, see), 4) *interpersonal focus* (categories: first person singular, second person plural, third person plural, indefinite pronoun), 5) *temporal references* (categories: future tense, past tense, present tense), 6) *lexical density and awareness* (categories: adverbs, verbs, article, exclusive, inclusive, preposition, quantifier), and 7) *personal and social concerns* (categories: achievement, bio, body, death, health, sexual, home, money, religion, family, friends, humans, social). For each cluster, I measure the normalized occurrences of each

Table 7.5: Summary of behavioral deviations in post-enrollment compared to expected (or predicted) behavior per cluster in terms of SMAPE, paired t-tests, and effect size (Cohen's d). Statistical significance reported as p-value, *<0.05, **<0.01, ***<0.001. Positive t or d indicates higher values in actual time series compared to the predicted time series.

Cluster	Model	100-Days			2-Weeks		
	SMAPE	SMAPE	t-test	d	SMAPE	t-test	d
Posting Behavior							
Average Daily Number of Posts							
C_0	11.09	24.45	-0.05	-0.01	30.73	-4.31 ***	-1.59
C_1	4.42	14.85	1.52	0.21	17.82	-3.68 ***	-1.35
C_2	5.78	17.45	3.49 ***	0.49	19.77	3.93 ***	1.44
C_3	4.20	18.00	5.76 ***	0.82	27.1	4.99 ***	1.84
C_4	5.85	16.25	2.02 *	0.29	17.04	1.07	0.39
Average Daily Number of Words							
C_0	22.69	42.85	1.10	0.16	54.22	-0.97	-0.36
C_1	11.24	24.99	-1.44	-0.2	19.46	-0.38	-0.14
C_2	11.31	24.5	2.60 *	0.37	23.28	1.46	0.54
C_3	6.40	17.86	3.03 ***	0.43	18.15	1.96 *	0.72
C_4	13.65	25.37	-0.18	-0.03	34.4	-3.72 ***	-1.37
Engagement Received							
Average Daily Number of Comments Received							
C_0	39.29	51.71	1.16	0.16	56.14	-0.41	-0.15
C_1	11.20	30.22	-0.35	-0.05	27.45	-0.41	-0.15
C_2	18.61	33.23	0.26	0.04	32.51	2.57 *	0.94
C_3	9.08	24.93	2.78 *	0.39	18.18	-0.18-	-0.07
C_4	25.63	31.38	0.46	0.07	26.55	-2.13 *	-0.78
Average Daily Number of Likes Received							
C_0	25.59	41.90	1.64	0.23	52.27	1.10	0.40
C_1	10.74	28.27	0.68	0.1	29.33	-0.97	-0.36
C_2	17.66	31.37	3.01 ***	0.43	33.87	3.10 ***	1.14
C_3	8.04	18.43	4.74 ***	0.67	17.45	0.72	0.26
C_4	13.74	28.2	1.70 *	0.24	32.34	-1.28	-0.47

LIWC category, and then compare the differences in the psycholinguistic use pre- and post-enrollment using independent sample t-tests.

7.4 RQ1: Findings on Observer Effect in Social Media Behavior

7.4.1 Deviation in Behavior

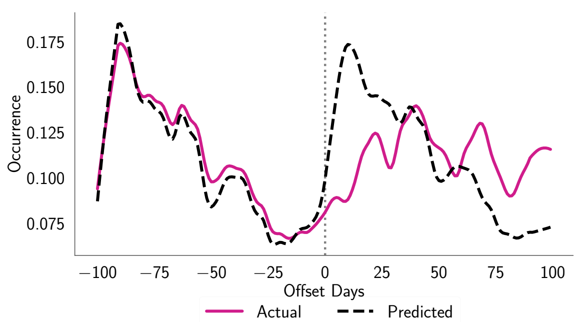
I calculate the deviation in the actual post-enrollment behavioral measures from predicted measures using autoregressive moving average (ARIMA) models that account for trends and seasonalities in time series. Table 7.5 summarizes the model metrics and observations with respect to changes in participants' social media behaviors.

Changes in Posting Behavior

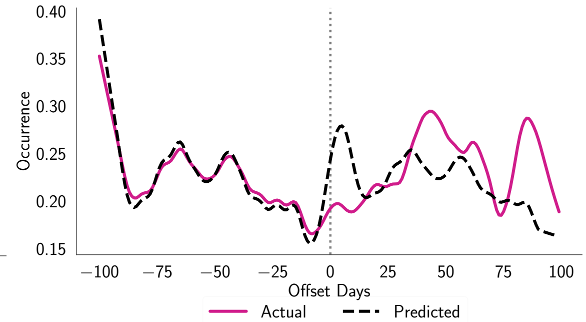
First, I note that the predictive models perform decently with models predicting the number of posts and number of words show mean SMAPEs of 6.27% and 13.05% respectively. However, the deviation in the post-enrollment data between predicted and actual values is higher. Looking at 100-days post-enrollment data, Clusters $\texttt{C}_2$ and $\texttt{C}_3$ show statistically significant deviations in both the measures of quantity and verbosity of posts, i.e., they post significantly more frequently and longer than their expected behavior. Next, focusing on the initial first two weeks post-enrollment, $\texttt{C}_2$ and $\texttt{C}_3$ show similar increase in posting. Interestingly, $\texttt{C}_0$ and $\texttt{C}_1$ show lower frequency of posting in the first two weeks, however, their posting behavior is closer to their expected posting behavior following the initial two weeks period. $\texttt{C}_4$ generally shows no significant change in behavior, except that these individuals tend to express shorter than expected posts in the first two weeks. Figure 7.7 show cluster-wise deviations in actual and expected time series of number of posts.

Changes in Engagement Received

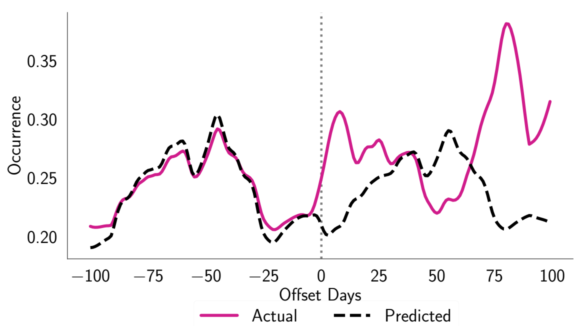
The models predicting engagement received, perform poorer than the above; models predicting number of comments and likes show mean SMAPEs of 20.76 and 15.15 respectively. Considering the 100-days of post-enrollment period, $\texttt{C}_3$ received higher than expected likes and comments and $\texttt{C}_2$ received higher than expected comments. The received engagements are also likely a correlate of these individuals' higher posting behavior as noted above. Looking at two-weeks' deviations, the findings suggest that $\texttt{C}_2$'s posts received immediate higher quantity of comments and likes, and $\texttt{C}_4$ received lower quantity of comments. Both $\texttt{C}_0$ and $\texttt{C}_1$ did not receive any significant deviations in the engagements received. Figure 7.8 shows an example time series plots of how the number of likes received evolved per cluster.



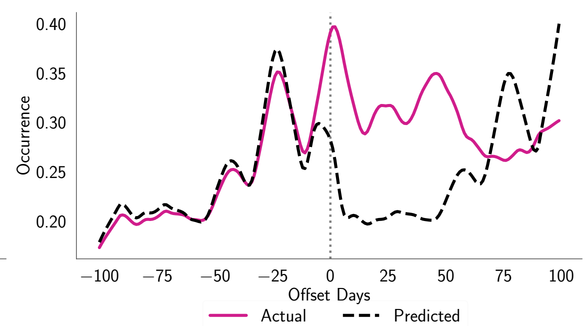
(a) Cluster $\mathtt{C}_0$



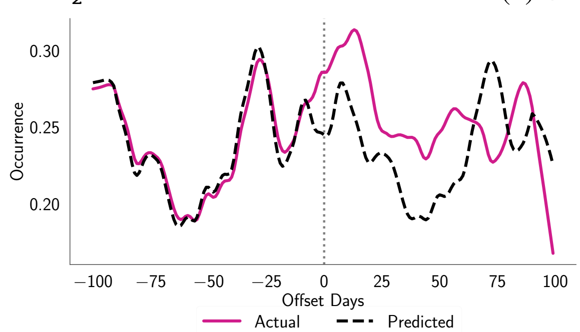
(b) Cluster $\mathtt{C}_1$



(c) Cluster $\mathtt{C}_2$

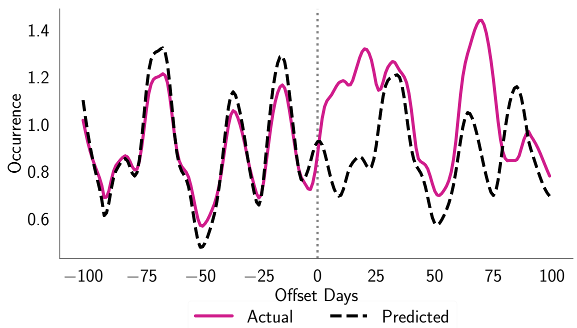


(d) Cluster $\mathtt{C}_3$

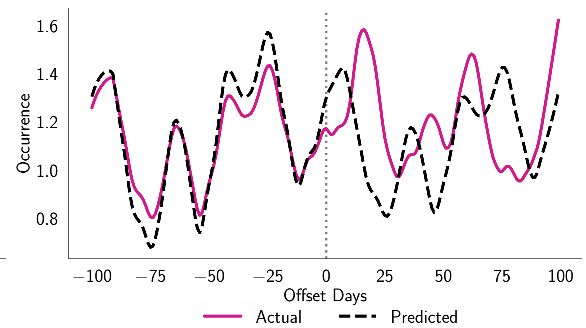


(e) Cluster $\mathtt{C}_4$

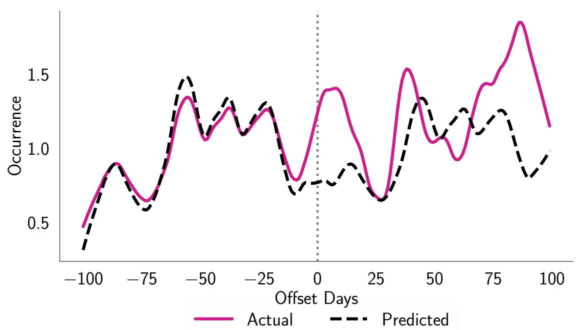
Figure 7.7: Evolution of the daily average number of posts per cluster in 100-days pre- and post- enrollment period. The dotted line in the center of each plot represents the date of enrollment (day 0).



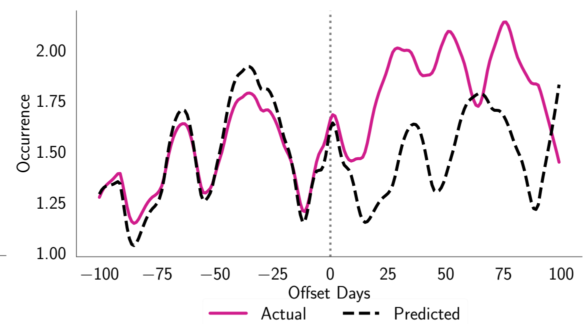
(a) Cluster $\mathtt{C}_0$



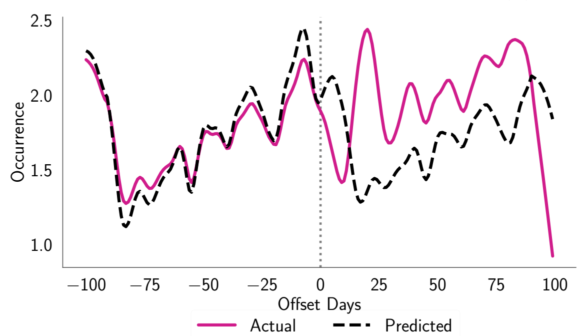
(b) Cluster $\mathtt{C}_1$



(c) Cluster $\mathtt{C}_2$



(d) Cluster $\mathtt{C}_3$



(e) Cluster $\mathtt{C}_4$

Figure 7.8: Evolution of the daily average number of likes per cluster in 100-days pre- and post- enrollment period. The dotted line in the center of each plot represents the date of enrollment (day 0).

Placebo Tests

To rule out the notion that the observed effects are at chance and not specific around participants' enrollment to the study, I conduct placebo tests. I repeat the time-series experiments on several (150) randomly permuted "placebo" enrollment dates in the pre-enrollment data of the participants. I measure the statistical significance as per t-test in the deviation in actual and predicted time series data for each of the permuted date for each cluster. Out of 150 permutations, $\mathtt{C}_0$ and $\mathtt{C}_4$ show significance in 2 and 1 instances respectively, and the other three clusters show no significant instances. Therefore, the probability of significant instance is close to 0 for all the clusters, revealing that the significance observed around the *actual enrollment dates* is not by chance.

7.4.2 Changes in Topical Themes

Table 7.6 summarizes the relative change in topical prevalence from pre- to post- enrollment for the participant clusters. Cluster $\mathtt{C}_0$ shows an increase in several themes of topics across travel, food, and news — all of which could be considered to be primarily more public content. These individuals also show an increased sharing about family gatherings, but decreased sharing about sports and celebratory events. Next, Cluster $\mathtt{C}_1$ shows increased sharing about holiday plans, family gatherings, and celebratory events, but decreased sharing about news-related content. Cluster $\mathtt{C}_2$ shows the least changes in expressiveness of content, with only decrease in sharing about food and social events. Cluster $\mathtt{C}_3$ shows varied changes, with increase in sharing about travel, food, and sports related content, whereas a decrease in more personal content such as holiday plans, work-life balance, family, and celebratory events. Finally, Cluster $\mathtt{C}_4$ shows an increase in sharing about food and family gatherings, whereas a decrease in holiday plans, news, and interests-related content.

Table 7.6: Changes in topical prevalence post-enrollment in the study. Statistical significance is computed as per independent-sample t-tests (* p<0.05, ** p<0.01, *** p<0.001). Significant values are shaded in blue for positive changes, i.e., higher average value in post-enrollment, and red for negative changes, i.e., lower average value in post-enrollment period.

Topic	% Change in Cluster				
	C_0	C_1	C_2	C_3	C_4
Travel & Locations	28.38 ***	-0.98	-7.69	25.14 *	-1.94
Food & Drinks	37.16 ***	2.28	-13.85 **	3.39 *	14.20 **
Holiday Plans	18.22 *	18.65 *	-7.22	-12.10 *	-10.21 *
News & Information	33.89 **	-14.25 ***	-6.19	-19.29 ***	-17.36 *
Work-Life Balance	-0.05	1.28	-8.93	-8.30 *	0.88
Family Gathering	56.72 ***	11.99 *	-7.43	3.41	36.54 ***
Social & Sports	-29.13 **	12.21	-4.54 *	66.77 ***	-14.62
Greetings & Celebrations	-11.58 ***	23.62 ***	-7.64	-28.64 ***	18.51
Friends & Family	-1.37	-5.98	-10.48	-12.79 ***	-9.33
Activities & Interests	-0.38	6.10	-21.42	-16.39	-30.58 **

7.4.3 Changes in Psycholinguistic Use

Finally, I examine the psycholinguistic changes in the clusters of participants. Table 7.7 shows the changes in psycholinguistic use.

First, the individuals in Cluster C_0 do not show any significant change in affective expressions except in the case of anger. In cognitive expressions, they show an increase in words related to certainty. In perception, *feel* and *see* decrease, whereas hear increases. They also show a decrease in first person singular pronoun use but increase in first personal plural pronoun use. Together, the pronoun use may indicate a decrease in self-attentional focus and increase in collective identity based language [131]. We also find a decrease in several function words, including adverbs, verbs, auxiliary verbs, quantifiers, and relatives. Among personal and social concerns, these individuals show an increase in achievement, home, and religion.

The individuals in Cluster C_1 do not show any significant change in affective, cognitive, and perceptive expressions. Among function words, they show a decrease in second person pronouns, and an increase in conjunction and inclusive. They also show a significant increase in social words, including the categories of family, friends, home, and religion. This also aligns with their topical changes in language post-enrollment. Therefore, these individuals do not show significant changes in non-content words, but significantly change their use of content words, i.e., we could assume that these individuals do not significantly change "how" they write, but do significantly change "what" they write.

The individuals in Cluster $\mathtt{C}_2$ show a significant decrease in a majority of affective, cognitive, and perceptive expressions, including anger, anxiety, negative affect, positive affect, causation, certainty, cognitive mechanics, inhibition, percept, and see. They show a decrease in the use of first person pronouns. In other function words, they show a decrease in past and present tense, article, verbs, inclusive, preposition, and relative. Again, in personal and social concerns, they show a decrease in friends, family, and home. Together, these psycholinguistic changes indicate that these individuals inhibit sharing about personal and self-expressive content, or prefer to share more about public-facing and less subjective content. This could be a sign of self-regulation among these individuals.

As above, the individuals in $\mathtt{C}_3$ show a significant decrease in several affective, cognitive, and perceptive attributes. They also decrease the use of first person singular pronouns, indicating lowered self-attentional focus, however, the use of third person pronouns significantly increase. Again, several function words decrease, including adverbs, verbs, and prepositions. In contrast to $\mathtt{C}_2$, $\mathtt{C}_3$ not only shows decreased negative affect and swear words but also increased positive affect and inclusive keywords. We also find an an increase in several social words, including family, humans, and social. These could be a manifestation of them wanting to self-present in a more socially desirable or positive way. Further, the decrease in *work* keywords might suggest that they were aware to not share work-related events on social media, particularly given the context that our study recruitment happened in workplace context.

Finally, the individuals in $\mathtt{C}_4$ show an increase in multiple affective expressions, including anger, negative affect, and swear, whereas a decrease in positive affect. Most cognitive and perceptive categories do not change, except a significant decrease in negation and feel. These

Table 7.7: Independent-sample t-tests in pre- and post- enrollment psycholinguistic (LIWC) use per cluster (* p<0.05, ** p<0.01, *** p<0.001). Significant values are shaded in blue for positive changes, i.e., higher average occurrence in post-enrollment, and red for negative changes, i.e., lower average occurrence in post-enrollment period.

LIWC	t-test				
	C_0	C_1	C_2	C_3	C_4
Affect					
Anger	2.01	-1.07	-2.02	-1.02	4.00 ***
Anxiety	1.41	0.03	-2.27	-1.957	1.93
Neg. Affect	0.83	-0.93	-2.60 **	-2.834 **	2.09
Pos. Affect	0.08	1.18	-4.49 ***	1.30 *	-2.06
Sadness	1.427	-0.42	1.52	-1.46	-0.61
Swear	1.134	0.60	-0.12	-3.06 **	7.53 ***
Cognition					
Causation	0.234	0.87	-2.69 **	-1.97	0.20
Certainty	4.08 ***	1.91	-2.12	-1.11	0.28
Cog. Mech.	1.32	0.86	-3.80 ***	-0.80	-0.93
Inhibition	-1.13	-1.37	-3.53 ***	-0.02	0.60
Discrepancies	-1.20	-1.61	1.08	-0.05	-0.55
Tentativeness	0.43	-1.17	1.79	-1.83	1.23
Feel	-2.31	0.87	-1.66	-3.12 **	-2.51
Hear	5.48 ***	0.50	2.39	1.27	1.41
Insight	-1.23	-0.141	-0.32	-2.39	0.90
Percept	-0.07	0.35	-4.74 ***	-1.23	-1.50
See	-2.31	-0.80	-4.77 ***	-1.41	-0.80
Interpersonal Focus					
1st P. Sing.	-7.29 ***	-1.00	-5.78 ***	-2.35	-4.17 ***
1st P. Plu.	2.25 *	0.47	-2.34	1.86	1.31
2nd P.	-1.43	-3.32 ***	5.71 ***	1.16	-0.70
3rd P.	-0.12	-0.63	-0.26	4.61 ***	-0.03
Indef. Pron.	-3.29 **	-1.33	-1.30	-3.43 ***	0.92
Fut. Tense	0.32	-0.65	2.32	-0.69	-0.36
Past Tense	1.85	0.28	-1.99	-0.158	2.61 **
Prs. Tense	-5.54 ***	0.19	-3.15 **	-6.49 ***	-1.90

LIWC	t-test				
	C_0	C_1	C_2	C_3	C_4
Lexical Density and Awareness					
Adverb	-3.00 **	-0.44	0.61	-3.36 ***	-2.46
Article	0.10	1.94	-3.60 ***	0.27	-1.34
Verb	-4.78 ***	0.47	-2.77 **	-5.53 ***	-1.36
Aux. Verb	-4.61 ***	0.40	0.10	-7.00 ***	-1.30
Conjun.	1.82	2.43	3.01 **	0.88	-0.15
Exclusive	1.05	-1.56	0.80	-1.92	-0.33
Inclusive	2.17	2.99 **	-3.47 ***	3.32 ***	-1.53
Negation	-1.38	-1.09	-0.90	-4.73 ***	-2.68 **
Preposition	1.47	-1.01	-3.27 **	-3.11 **	-2.28
Quantifier	-2.34	0.96	0.71	-0.06	0.50
Relative	-2.20	-1.16	-3.65 ***	-1.57	-2.98 **
Personal and Social Concerns					
Achvmt.	3.28 **	-0.91	-2.61**	0.14	-1.08
Bio	1.57	2.57	0.09	-0.20	-2.77 **
Body	-1.72	0.74	1.03	-1.34	-1.73
Death	0.43	1.99	-0.931	-0.162	-0.45
Family	1.14	2.64 **	-2.06	3.66 ***	0.29
Friends	-2.08	0.35 *	-1.46 *	-2.01	-1.17
Health	0.52	0.47	-0.34	-0.11	-0.81
Home	3.14 **	2.77 **	-2.19	1.39	0.08
Humans	-2.94 **	-1.67	0.51	2.85 **	-0.62
Money	-1.81	-1.06	-1.05	1.01	-1.24
Religion	2.29	2.48	-1.06	-0.87	-0.14
Sexual	1.62	-0.27	1.07	-0.62	0.56
Social	-1.02	-0.63	-1.542	2.79 **	0.30
Work	0.29	-0.58	-4.57 ***	-2.96 **	-1.74

individuals also show a decreased use of first person singular pronouns, but an increase in past tense. Most other function words and social words do not show significant change, except significant reduction in the use of adverbs, preposition, relative, and bio.

7.5 RQ2: Explaining Observer Effect Based on Individual Differences

This section targets the second research question to explain observer effect on social media behavior, through theories relating to individual differences and psychological traits. For each cluster, I examine the offline (psychological) characteristics, and evaluate the behavioral and linguistic changes as observed in the social media behavior. I contextualize and interpret the findings by drawing upon the literature in human behavior, psychology, and social science (as discussed in section 7.1). Table 7.8 summarizes the observations from this study.

Cluster $\mathtt{C}_0$ contains individuals with high levels of conscientiousness and sleep quality, and low levels of openness and cognitive ability, suggesting that these individuals are more likely to be routine-oriented and pragmatic. Prior literature notes that high conscientiousness is associated with self-monitoring [571]. This could be associated with the changes in their posting behavior: they tend to significantly reduce their posting immediately after enrollment, which however, gets back closer to their normalcy over time. This aligns with behavioral amendments as a form of *habituation* explained in behavioral science [117]. Linguistically, these individuals show an increased sharing about public-facing information, and when coupled with the observation of decreased first person singular pronouns, can be considered to be reduced self-attentional focus, and increased sharing about events attended as a part of group.

Cluster $\mathtt{C}_1$ contains individuals with high cognitive ability and low neuroticism, suggesting that they are likely to be reasonable and composed in day-to-day and general aspects of life [42]. While their posting was significantly decreased in the first two weeks, posting behavior became closer to the normal subsequently. These individuals show an increase in **sociality** after enrollment [204]. One possible explanation of their behavior could be based upon Middleton, Buboltz, and Sopon's observation that individuals with higher cognitive ability are less likely to show psychological reactance [416]. Likewise, observations on $\mathtt{C}_1$ also aligns with prior work where the increased use of family-related keywords are known to be associated with lower self-monitoring skills [290]. Together, these individuals might have lower self-monitoring skills, be less bothered by the aspect of being "observed", and be comfortable to continue sharing their social and personal life on social media.

Cluster $\mathtt{C}_2$ consists of individuals with high levels of neuroticism, cognitive ability, negative affect, and anxiety, and low levels of extraversion, agreeableness, conscientiousness, positive affect, and sleep quality. These characteristics suggest that these individuals are likely to be more withdrawn and more prone to stress and irritability [42]. These individuals show decreased sharing about social topics such as food and drinks, and sports and social events. This is also reflected in their psycholinguistic use of lowered personal and social words such as family, friends, and home. However, these individuals increased their posting activity post-enrollment. Their higher volume of post-enrollment posting behavior could be associated with higher self-monitoring skills as per prior work [287]. These individuals also received greater engagement in terms of likes and comments — this could be a function of heightened information seeking on social media, which is known to be associated with higher neuroticism [569], as also in the case of $\mathtt{C}_2$.

Cluster $\mathtt{C}_3$ consists of individuals with high extraversion, agreeableness, and conscientiousness. Extraversion is known to positively correlate with public self-consciousness [596] and self-monitoring [43]. Similar to $\mathtt{C}_2$, greater posting behavior in $\mathtt{C}_3$ could be a manifestation of high self-monitoring skills [287]. Further, high conscientiousness could also dictate a desire to appear as "good" participants or self-present in a more desirable way [42] — this could be reflected in their increased social media activities, increased positive affect, and decreased negative affect and swear words. Then, high agreeableness is known to be associated with people's likelihood to seek acceptance and maintaining social connections [569]. Similar phenomenon is observable in our findings as these individuals posts elicited greater number of likes and comments, compared to before enrolling in the study. Further, prior work situated theory of planned behavior in explaining greater intention to post online for individuals with high agreeableness and extraversion [480].

Cluster $\mathtt{C}_4$ consists of individuals with high openness. Although their posting does not significantly change immediately, they show a significant increase in their posting behavior throughout the post-enrollment period, compared to expected behavior. They also show significant linguistic changes in this period. In particular, they show increased sharing about many personal and social aspects of life, despite a significant reduction in first person singular pronouns and many function words. At a meta-level, these individuals show lowered

Table 7.8: Summary of Findings.

Cl.	Traits	Behavior	Topics	Psycholinguistics	Notes / Descriptor
c_0	High (Conscientiousness, Sleep Quality), Low (Openness, Cognitive Ability)	Posting significantly reduces in the initial few days, then back to normalcy (Figure 7.7a)	Increased sharing about public-facing information (Table 7.6)	Increased (anger, achievement, home, religion), Decreased (feel, first person singular, present tense, function words, friends, humans) (Table 7.7)	High conscientiousness is associated with self-monitoring. Habituation in posting behavior. Decreased self-attentional focus.
c_1	High (Cognitive Ability), Low (Neuroticism)	Posting significantly decreased in the first two weeks, then back towards normalcy (Figure 7.7b)	Increased sharing about family gathering, social, and online greeting related activities (Table 7.6)	Increased (social words), Decreased (2nd person) (Table 7.7)	These participants are trait-wise more reasonable and composed. They show high sociality post-enrollment. Low psychological reactance and low self-monitoring skills; less bothered about being "observed".
c_2	Low (Extraversion, Agreeableness, Conscientiousness, PA, Sleep Quality), High (Neuroticism, Cognitive Ability, NA, Anxiety)	Posting significantly increased throughout. Greater engagement received. (Figure 7.7c)	Decreased sharing about food and social topics (Table 7.6)	Increased (hear, future tense), Decreased (affective, cognitive, perceptive, 1st person pronouns, function words, social words) (Table 7.7)	Trait-wise, they may be more withdrawn, and prone to stress and irritability. High self-monitoring skills, and heightened information seeking (associated with high neuroticism).
c_3	High (Extraversion, Agreeableness, Conscientiousness, PA, Sleep Quality), Low (Neuroticism, NA, Anxiety)	Posting significantly increases throughout. Greater engagement received. (Figure 7.7d)	Decreased sharing about personal events (Table 7.6)	Increased (social words, third person pronouns), Decreased (affective, cognitive, perceptive, first person pronouns, function words) (Table 7.7)	Desire to self-present in a more desirable way. Likelihood to seek acceptance and maintain social connections.
c_4	High Openness	No immediate significant difference in posting frequency, but posting significantly increases throughout. More likes received. (Figure 7.7e)	Decreased sharing about news and holiday plans. Increased sharing about food/family gathering (Table 7.6)	Increased (anger, NA, swear, past tense), Decreased (PA, negation, feel, 1st person singular, function words) (Table 7.7)	Self-regulation. Less personal-content. High psychological reactance, manifested in detached sharing about personal content.

use of negations and exclusives, suggesting lowered cognitive complexity in language — which could be associated with less personal content [475]. These changes may suggest that these individuals are likely to self-regulate their social media behaviors to present selective aspects of life without sharing too intimate content. Again, greater openness is known to be associated with high psychological reactance [568], which could be manifested in detached sharing about personal and first person singular content. Prior work has associated openness with greater resiliency and externally induced behavioral changes [420], however, the interplay of such a characteristic with observer effect remains to be examined further.

7.6 Discussion

7.6.1 Theoretical Implications

Correcting biases in prospective use of social media as a wellbeing sensor

This study provides insights regarding the prevalence and degree of observer effect in the social media behavior. In particular, I examined how people deviated from their normative (or expected) behaviors after enrolling in a multimodal sensing study. This study informs research about how to correct for data, biases, and models when implementing practical and prospective data-driven assessments and interventions. In this regard, this study contributes to the recommendations made by Ruths and Pfeffer in correcting biases of big-data technologies [525]. The study findings help us to gauge what to expect when social media is used as passive sensor in prospective setting. This would help us be more cognizant about which individuals might significantly deviate from their otherwise normative behaviors, and accordingly build personalized models that are robust to people's baseline traits and tendencies to be impacted by observer effect.

Generating hypotheses

The findings of this study can help to generate hypotheses relating to observer effect in social media and multimodal sensing. These hypotheses can be tested and evaluated individually and rigorously. RQ2 explained the findings through theories in psychology and social science literature. These associations can be formulated as testable hypotheses in future research. This study also motivates to incorporate other intrinsic and social processes such as self-censorship and privacy perceptions, which may also interact with social media behavioral change [151, 403]. It is also important to note how the findings are also an artifact of the domain and the participant pool. This study is conducted on a specific participant pool of information workers in the context of workplace settings. Such a factor may have an effect on the changes observed in work-related language (in Table 7.6 and Table 7.7). Future experiments can explore more conclusive and generalizable evidences about observer effect, and whether these are opportunities or challenges in other situations and contexts.

Complementary Assessments of Observer Effect

As already noted before, observer effect has not only been hard to study, but also there is no established gold-standard of measuring observer effect. In this regard, this is the first study of measuring observer effect on social media behavior. Due to the lack of direct means to measure success and construct validity of this research, I tested and situated the findings with existing theories. While this study approached targeted to draw passive and more objective form of assessment, it is also important to account for self-reported assessments and perceptions about observer effect. Therefore, this work motivates to design and conduct surveys and interviews, which would help us gauge complementary information about how observer effect manifests in social media behavior.

Self-selection and "Who is the observer?"

In this particular scenario, observers were a set of researchers to whom the participants willingly shared their data, and there was a data sharing protocol in place. Again, the participants self-selected themselves in the study in return of a participation compensation. However, observer effect can manifest in many other cases in different combinations of other kinds of observers or data sharing terms. Therefore, understanding the role of these factors with respect to observer effect would be important to measure and correct for observer effect as needed in real-world situations.

7.6.2 Implications for Researchers and Practitioners

This research showed that individuals are likely to deviate from their expected behavior when subjected to real-time and prospective data collection settings — attributed as some form of "observer effect". Such an effect needs to be accounted for to successfully instrument real-time applications of human-centered social media based assessments. The computational approaches adopted in this study can be used to measure observer effect in various contexts. Researchers can use such approaches to identify cases of observer effect-based deviations, and build predictive models robust to such effects in a person-centric fashion. In addition, this study reflects that self-reported psychological traits can not only be used to stratify and cluster individuals, but also to explain their behavioral changes due to observer effect. Similar approaches can be used to build person-centered modeling and interventions for different groups of individuals.

Besides highlighting the potential assessment-centric biases, this study also motivates us to critically reflect and rethink the implications surrounding the individuals' autonomy in using technologies. Individuals primarily use social media platforms to share and connect with others. However, if external interventions potentially interfere with their social media use or make them feel uncomfortable or surveilled, then the fundamental goals and expectations of using social media platforms are interfered. Such an unintended consequence needs to be evaluated by researchers and practitioners while building digital data-driven assessments. Therfore, this work encourages us to critique the trade-offs of the harms and benefits of using social media based technologies for wellbeing and behavior assessments.

CHAPTER 8
CONCLUSION

I began this dissertation with an aim to evaluate and showcase social media as a viable passive sensor of wellbeing, particularly in situated communities — particularly, two communities with which we are likely to closely relate ourselves with, college campuses and workplaces. This dissertation showed multiple studies towards this goal. In conclusion, I note some of the major unique aspects of the work presented in this dissertation. First, this work is motivated by the social-ecological model [102] that human behaviors and wellbeing are embedded in the complex interplay of the individual, community, and society. To get a better understanding of wellbeing, we need to incorporate the situated context, and this dissertation focuses on situated communities as examples to consider situated contexts. Next, given that the targeted problems are in a real-world context, it is important to account for the confounds and latent factors that may impact an individual's wellbeing. While randomized experiments would have been ideal, such settings are often infeasible and unethical in several practical problem scenarios. In this regard, I developed computational and causal frameworks that minimize the confounding factors, drawing upon machine learning, natural language analysis, and statistical modeling techniques. The theory-driven computational methodologies proposed in this dissertation can be adapted in several related settings and to help build tailored and timely supportive interventions. Further, this dissertation proposed methodologies to combine social media with multimodal sensing, leveraging complementary strengths of multiple sensing modalities towards a comprehensive understanding of human behavior and wellbeing. Together, this dissertation propels the vision towards leveraging multimodal sensing to build human-centered technologies for wellbeing.

I recognize that the proposed approaches bear real-world implications, and it is important to introspect and interpret these online-data-driven offline inferences — this forms a cross-cutting theme in this dissertation. These studies bear implications to help inform and develop real-world supportive interventions. For example, stakeholders at college campuses or workplaces can build real-time dashboards and assessment techniques to proactively help the community members. However, a question that remains largely unexplored, is how would these systems perform in the real-world? That is, the in-practice utility and ecological validity of these systems remain largely unknown. Relatedly, this dissertation targeted the problem of "observer effect". I studied observer effect in social media behavior in the context of a multimodal sensing study. I provided a methodology to measure observer effect, and found insights about the prevalence and degree of observer effect, and how it varies with individual characteristics. This dissertation motivates future research in evaluating such questions on the in-practice utility of social media and multimodal sensing.

I illustrate some open questions in the problems discussed in this dissertation, and more generally in this problem space. These questions open up future directions to think, evaluate, and address. For example, there is a lack in understanding of ground-truth? What does it mean to be ground-truth and how do we collect that? Again, these assessments are not immune to biases, for example, social media data suffers from self-selection and self-censorship biases. Therefore, how do we minimize and address these biases going forward? How do we ensure that these methods also benefit beyond those who afford or use these technologies? Again, these assessments can be misused, and findings can be misinterpreted to reinforce existing societal biases. While it is important to build transparent and interpretable models without compromising with performance, we also need to ensure that the algorithms are not misused for unintended and unethical consequences. That is, we need to navigate through the harms and benefits of these algorithms and tools, and build rigorous but responsible and ethical technologies. These questions pose both challenges as well as opportunities to strive towards building better future technologies for wellbeing. It is also important to critically study these questions and bring together multi-stakeholder viewpoints to realize the research and practical impact. I envision these topics to encourage further discussions among researchers, ethicists, and technology users.

I end the dissertation with a few future directions to pursue in the area of research with respect to social media, wellbeing, and human-centered machine learning.

Developing Human-centered Approaches with Technology Tailored to a Person's Situations, Demands, and Needs. The popular phrase "one size does not fit all" applies in several scenarios, including computational techniques to assess individual and collective attributes. As also noted in this dissertation, all individuals are not the same and have different experiences, the between-person variability in data may impact predictions of an individual's underlying psychology, routines, and other personal attributes. Further, generalized models can be exclusionary, and be reinforcing stereotypes and existing societal biases. Such approaches may suffer from the limitations of oversimplifying the social reality, and may misrepresent or suppress the voices of the underrepresented and historically marginalized groups. Therefore, it is only imperative that we build algorithms catering to each individual's constraints and factors. For instance, incorporating additional offline context that captures factors explaining posting (or not posting) behavior on social media can boost the ability of social media to predict individual outcomes. By accounting for people's voices (as naturalistically expressed on social media), we can build technologies that incorporate individual autonomy and more individual-facing approaches by stakeholders. Such work will help build tools that quantify and capture wellbeing constructs on a continuous and real-time basis and can help identify the (often) unknown concerns of wellbeing.

Facilitating Online Technology Supported Remote Functioning and Wellbeing Future research can build upon this dissertation to contribute towards the theme of "Future of Work". In light of the ongoing COVID-19 pandemic, we can build social media and ubiquitous technologies that facilitate remote worker functioning [153, 541]. The pandemic has reinforced the importance of and dependence on computing technologies in our lives. Going forward, we can build technologies that leverage the advantages of virtual interactions, and mitigate the challenges due to the lack of physical and face-to-face interactions. This dissertation revealed how self-initiated online data such as individuals' self-presentation on LinkedIn to assess their role ambiguity [536], and company reviews on Glassdoor to assess the company culture [155], and similar approaches can be used to solve problems associated with the challenges of work and technology-assisted work with the changing scenarios. For instance, remote work settings can lead to overlapping boundaries of personal and professional lives and cause additional stress. Again, new disadvantages and disparities may have begun, e.g., defining workplace culture would now need to account for remote collaboration and remote dynamics. Further, online settings may bring in the complexities of online antisocial behaviors, such as harassment and discrimination.

Evaluating Prospective Utility and Designing Online Social Platforms for Wellbeing Support. This dissertation stressed how a significant body of research in this area relies on data that is retrospectively collected, and if we envision a future with social media technologies for wellbeing interventions, we need to ask about the efficacy of these algorithms perform in *prospective* data collection and assessments. This dissertation motivates us to build approaches that correct for biases in the in-practice utilization of social media data-driven assessments. Such evaluations would also provide insights to drive designing and building online wellbeing interventions for different populations. For instance, we found the efficacy of counseling interventions via social media on college students [544]. Future research can build upon the implications of such studies to design systems that facilitate peer mental health support among different communities. This body of work needs to be expanded through experimental studies of how certain interventions help (or do not help) individuals. We need to recognize the sensitivity of this work, and to negotiate the challenges, research should be conducted in close collaboration with mental health support volunteers, psychologists and clinicians, platform owners, and users.

Understanding Harms/Benefits of Computational and Data-Driven Assessments Social media and ubiquitous technologies serve a lot of opportunities and advantages in assessing many psychological, cognitive, and social outcomes. However, these come at a cost. Often, collecting and analyzing such data and using data-driven insights for real-world decisions can mean compromising with privacy and ethics. There are lingering questions about the circumstances under which such inferences should be made, and, what should be the best practices. Despite the best of intentions, these methodologies can lead to expectation mismatches, and individuals may perceive intrusiveness and dissatisfaction about such algorithmic inferences on data without their consent or awareness [213]. There are bad actors in the online and offline world, and they can potentially use these inferences in ethically questionable ways, leading to compromised privacy, defying expectations, and damaging trust between individuals and technology. We need to build approaches that balance the trade-offs of the risks and benefits. This direction of work can revisit and recommend guidelines towards transparency and accountability, and bear multi-stakeholder implications in human-centered computing research of wellbeing.

Appendices

APPENDIX A

MODELING ORGANIZATIONAL CULTURE

A.1 Detailed codebook

Table A.1: Summary of job aspects, descriptions, and corresponding measure in organizational culture.

Aspect	Validation Source
Social Skills	
Instructing: Teaching others how to do something	Human-Relations Model [495]
Service Orientation: Actively looking for ways to help people	Affiliative Norms [135]
Interests	
Social: Working with, communicating with, & teaching people. These occupations often involve helping or providing service to others.	Need for Security: Value [304]
Enterprising: Starting and carrying out projects. These occupations can involve leadership and decision making. Sometimes require risk taking and often deal with business.	Self-Actualizing Norms [135]
Conventional: Following set procedures & routines. These occupations include working with data & details instead of ideas. They offer a clear line of authority to follow.	Conventional Norms [135]
Work Values	
Achievement: Result oriented, allows strongest abilities to give a feeling of accomplishment. Corresponding needs: Ability Utilization and Achievement.	Process-Oriented vs Results-Oriented (Practice) [304]
Independence: Allow employees to work on their own and make decisions. Corresponding needs: Creativity, Responsibility and Autonomy.	Dependent Norms [135]
Recognition: Offer advancement, leadership potential, & often considered prestigious. Corresponding needs: Advancement, Authority, Recognition and Social Status.	Need for Security (Value) [304]
Relationships: Allow employees to help others & work with co-workers in a friendly environment. Corresponding needs: Co-workers, Moral Values & Social Service.	Affiliative Norms [135]
Support: Offer supportive management that stands behind employees. Corresponding needs are Company Policies, and Human Relations and Technical Supervision	Supervision [240]
Working Conditions: Offer job security & good working conditions. Corresponding needs: Activity, Compensation, Independence, Security, Variety, Working Conditions.	Need for Security (Value) [304]
Work Styles	
Achievement/Effort: Establishing and maintaining personally challenging achievement goals and exerting effort toward mastering tasks.,	Achievvment Norms [135]

Adaptability/Flexibility: Being open to change (positive or negative) and to considerable variety in the workplace.,	Open-Systems Model [495]
Concern for Others: Being sensitive to others' needs and feelings and being understanding and helpful on the job.,	Humanistic-Encouraging Norms [135]
Cooperation: Being pleasant with others on the job and displaying a good-natured, cooperative attitude.	Parochial vs Professional (Practice) [304]
Independence: Developing one's own ways of doing things, guiding oneself with little or no supervision, and depending on oneself to get things done.,	Self-Actualizing Norms [135]
Initiative: Willingness to take on responsibilities and challenges.,	Work Centrality (Value) [304]
Integrity: Honest and ethical conduct	Normative vs Pragmatic (Practice) [304]
Leadership: Willingness to lead, take charge, and offer opinions and direction	Involvement [240]
Self Control: Maintaining composure, keeping emotions in check, controlling anger, and avoiding aggressive behavior, even in very difficult situations.,	Process-Oriented vs Results-Oriented (Practice) [304]
Social Orientation: Preferring to work with others rather than alone, and being personally connected with others on the job.,	Humanistic-Encouraging Norms [135]
Stress Tolerance: Accepting criticism and dealing calmly and effectively with high stress situations.	Need for Security (Value)[304]
Work Activities: Interacting with Others	
Assisting and Caring for Others: Providing personal assistance, medical attention, emotional support, or other personal care to others (coworkers, customers, patients)	Affiliative Norms [135]
Coaching and Developing Others: Identifying developmental needs of others and coaching, mentoring, or otherwise helping others to improve their knowledge or skills	Humanistic-Encouraging Norms [135]
Developing and Building Teams: Encouraging and building mutual trust, respect, and cooperation among team members.,	Humanistic-Encouraging Norms [135]
Establishing & Maintaining Interpersonal Relationships: Developing constructive & co-operative working relationships with others, & maintaining them	Affiliative Norms [135]
Guiding, Directing, and Motivating Subordinates: Providing guidance & direction to subordinates, like setting performance standards & monitoring performance	Power Norms [135]
Monitoring and Controlling Resources: Monitoring and controlling resources and overseeing the spending of money	Loose-Control vs Tight-Control (Practice) [304]
Resolving Conflicts and Negotiating with Others: Handling complaints, settling disputes, and resolving grievances and conflicts, or otherwise negotiating with others	Need for Authority (Value) [304]
Training and Teaching Others: Identifying the educational needs of others, developing formal educational or training programs or classes, and teaching others	Human-Relations Model [495]
Work Context: Structural Job Characteristics	
Consequence of Error: Serious would the result usually be if the worker made a mistake that was not readily correctable,	Perfectionist Norms [135]

Freedom to Make Decisions: Much decision making freedom, without supervision, does the job offer,	Dependent Norms [135]
Frequency of Decision Making: Frequently is the worker required to make decisions that affect other people, the financial resources, and/or the image and reputation of the organization,	Employee-Oriented vs Job-Oriented (Practice) [304]
Importance of Being Exact or Accurate: Is being very exact or highly accurate in performing this job,	Perfectionist Norms [135]
Level of Competition: To what extent does this job require the worker to compete or to be aware of competitive pressures,	Competitive Norms [135]
Structured versus Unstructured Work: To what extent is this job structured for the worker, rather than allowing the worker to determine tasks, priorities, and goals,	Need for Security (Value) [304]
Work Schedules: Regular are the work schedules for this job	Loose-Control vs Tight-Control (Practice) [304]
Work Context: Interpersonal relationships	
Face-to-Face Discussions: Face-to-face discussions with individuals or teams,	Meetings [240]
Frequency of Conflict Situations: Conflict situations the employee has to face,	Need for Authority (Value) [304]
Responsibility for Outcomes and Results: responsible for work outcomes and results of other workers,	Avoidance Norms [135]
Work With Work Group or Team: Work with others in a group or team	Teamwork-Conflict [240]

APPENDIX B

PERSON-CENTERED CONTEXTUALIZATIONS

B.1 Detailed models

Table B.1: Generalized Models: Predicting psychological constructs with social media using the entire data of all participants. Prediction algorithms used include Ridge, Elastic Net (ElNet), Support Vector Regressor (SVR), XGBoost (XGB), Gradient Boosted Random Forest (GBR), and Multilayer Perceptron Regressor (MLP). Reported accuracy numbers are Symmetric Mean Absolute Percentage Error (SMAPE) and Pearson's correlation coefficient (r), which are pooled in k-fold cross-validation (k=5). The bold-faced number in each row indicate the best performing model for that construct.

Construct	Algorithm											
	Ridge		ElNet		SVR		XGB		GBR		MLP	
	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE
Cognitive Ability												
Shipley (Abstraction)	**0.25**	**6.81**	-0.19	6.78	0.22	6.75	0.18	6.85	0.23	6.78	0.09	17.75
Shipley (Vocabulary)	**0.29**	**4.13**	-0.14	4.33	0.24	4.14	0.18	4.28	0.22	4.24	0.14	10.04
Personality Traits												
Openness	**0.25**	**6.89**	-0.13	6.65	0.1	6.60	0.15	6.68	0.15	6.71	0.12	13.04
Conscientiousness	**0.13**	**7.29**	-0.14	7.07	0.08	7.02	0.04	7.34	0.06	7.28	0.04	11.35
Extraversion	0.13	8.93	-0.14	8.69	-0.06	8.69	0.17	8.61	**0.17**	**8.54**	0.14	12.54
Agreeableness	**0.17**	**5.84**	-0.15	5.78	-0.04	5.76	0.18	5.89	0.16	6.09	0.12	11.9
Neuroticism	**0.12**	**13.56**	-0.17	13.17	-0.14	13.11	0.05	13.59	0.06	13.37	0.09	13.87
Affect and Wellbeing												
Pos. Affect	0.07	7.27	-0.07	6.88	0.11	6.83	**0.13**	**7.10**	0.13	6.92	0.07	12.81
Neg. Affect	**0.11**	**10.90**	-0.17	10.89	-0.11	10.89	-0.05	11.51	-0.04	11.66	-0.0	20.08
Anxiety (STAI)	**0.12**	**9.66**	-0.14	9.66	-0.1	9.54	-0.06	10.2	-0.02	9.97	0.07	14.85
Sleep Quality (PSQI)	**0.15**	**16.02**	-0.14	15.52	-0.12	15.15	0.17	15.17	0.16	15.28	0.08	23.01

Table B.2: **Contextualized Models**: Predicting psychological constructs with social media separately for each behaviorally contextualized clusters. Prediction algorithms used include Ridge, Elastic Net (ElNet), Support Vector Regressor (SVR), XGBoost (XGB), Gradient Boosted Random Forest (GBR), and Multilayer Perceptron Regressor (MLP). Reported accuracy numbers are Symmetric Mean Absolute Percentage Error (SMAPE) and Pearson's correlation coefficient (r), which are pooled in k-fold cross-validation (k=5). The bold-faced number in each row indicate the best performing model for that construct.

Construct	**Algorithm**											
	Ridge		**ElNet**		**SVR**		**XGB**		**GBR**		**MLP**	
	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE
Cognitive Ability												
Shipley (Abstraction)	**0.23**	**6.88**	-0.17	6.82	-0.09	6.8	0.22	6.77	0.19	6.84	0.09	19.56
Shipley (Vocabulary)	**0.21**	**4.25**	0.03	4.33	0.03	4.16	0.21	4.25	0.28	4.11	0.13	15.82
Personality Traits												
Openness	**0.29**	**6.08**	0.01	6.66	0.07	6.6	0.15	6.75	0.13	6.81	0.08	22.3
Conscientiousness	**0.16**	**7.08**	-0.1	7.1	-0.0	7.04	0.06	7.25	0.08	7.24	-0.02	22.64
Extraversion	**0.21**	**8.46**	-0.07	8.72	-0.06	8.71	0.16	8.66	0.16	8.62	0.1	22.66
Agreeableness	**0.19**	**5.89**	-0.14	5.8	-0.06	5.78	0.1	6.14	0.13	6.0	0.02	21.74
Neuroticism	0.14	13.22	-0.15	13.22	-0.1	13.19	0.12	13.37	**0.18**	**13.09**	0.01	20.72
Affect and Wellbeing												
Pos. Affect	**0.14**	**6.90**	-0.01	6.88	0.04	6.82	0.07	7.31	0.04	7.35	0.06	15.25
Neg. Affect	**0.13**	**10.89**	0.01	10.87	0.0	10.8	0.03	11.26	0.01	11.36	0.01	22.23
Anxiety (STAI)	**0.21**	**8.51**	-0.04	9.68	-0.13	9.55	0.01	10.11	0.06	9.81	0.07	16.81
Sleep Quality (PSQI)	0.21	10.06	-0.05	15.49	-0.01	15.12	0.20	11.43	**0.25**	**10.59**	0.07	27.12

Table B.3: Generalized Models with PCA: Predicting psychological constructs with social media using the entire data of all participants, after applying PCA-transformed features. Prediction algorithms used include Ridge, Elastic Net (ElNet), Support Vector Regressor (SVR), XGBoost (XGB), Gradient Boosted Random Forest (GBR), and Multilayer Perceptron Regressor (MLP). Reported accuracy numbers are Symmetric Mean Absolute Percentage Error (SMAPE) and Pearson's correlation coefficient (r), which are pooled in k-fold cross-validation (k=5). The bold-faced number in each row indicate the best performing model for that construct.

Construct	**Algorithm**											
	Ridge		**ElNet**		**SVR**		**XGB**		**GBR**		**MLP**	
	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE
Cognitive Ability												
Shipley (Abstraction)	0.10	11.95	-0.19	6.78	**0.22**	**6.68**	0.21	6.86	0.2	6.95	-0.06	22.44
Shipley (Vocabulary)	0.11	7.57	-0.09	4.32	**0.23**	**4.10**	0.22	4.4	0.22	4.42	0.12	29.34
Personality Traits												
Openness	0.12	10.89	-0.13	6.65	**0.23**	**6.40**	0.16	6.68	0.16	6.64	0.14	15.87
Conscientiousness	0.03	11.25	-0.14	7.07	**0.13**	**6.97**	0.12	7.13	0.11	7.19	0.08	16.22
Extraversion	0.11	13.43	-0.14	8.69	**0.20**	**8.47**	0.17	8.73	0.19	8.66	0.1	16.25
Agreeableness	0.11	10.1	-0.15	5.78	**0.20**	**5.66**	0.11	6.01	0.12	5.98	0.11	14.02
Neuroticism	0.02	22.47	-0.17	13.17	-0.02	13.29	0.07	13.54	0.05	13.48	**0.09**	**21.8**
Affect and Wellbeing												
Pos. Affect	0.05	11.38	-0.07	6.88	**0.13**	**6.79**	0.09	7.15	0.03	7.24	0.03	29.89
Neg. Affect	0.07	17.54	-0.17	10.89	-0.07	10.82	**0.10**	**11.29**	0.10	11.38	0.12	22.91
Anxiety (STAI)	**0.07**	**15.41**	-0.12	9.66	-0.0	9.51	0.05	9.93	0.0	10.06	0.10	30.92
Sleep Quality (PSQI)	0.04	26.74	-0.14	15.52	0.10	15.04	0.09	15.94	**0.12**	**15.68**	0.12	24.60

Table B.4: Contextualized Models with PCA: Predicting psychological constructs with social media separately for each behaviorally contextualized clusters, after applying PCA-transformed features. Prediction algorithms used include Ridge, Elastic Net (ElNet), Support Vector Regressor (SVR), XGBoost (XGB), Gradient Boosted Random Forest (GBR), and Multilayer Perceptron Regressor (MLP). Reported accuracy numbers are Symmetric Mean Absolute Percentage Error (SMAPE) and Pearson's correlation coefficient (r), which are pooled in k-fold cross-validation (k=5). The bold-faced number in each row indicate the best performing model for that construct.

Construct	Algorithm											
	Ridge		ElNet		SVR		XGB		GBR		MLP	
	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE
Cognitive Ability												
Shipley (Abstraction)	0.10	12.78	-0.05	6.91	0.14	6.75	**0.21**	**6.97**	0.16	7.11	-0.08	40.14
Shipley (Vocabulary)	0.03	8.74	0.09	4.31	0.14	4.14	0.17	4.49	**0.21**	**4.46**	0.06	51.14
Personality Traits												
Openness	0.14	12.16	-0.0	6.68	**0.24**	**6.45**	0.09	7.02	0.16	6.96	0.05	26.83
Conscientiousness	0.08	12.34	-0.06	7.10	**0.14**	**6.94**	0.05	7.53	0.03	7.6	0.06	27.91
Extraversion	0.06	14.02	0.02	8.69	0.16	8.63	**0.21**	**8.69**	0.24	8.54	0.11	26.64
Agreeableness	0.08	10.79	-0.15	5.8	**0.13**	**5.75**	0.05	6.23	0.06	6.24	0.09	27.29
Neuroticism	0.04	23.19	-0.12	13.31	0.05	13.17	0.04	13.66	0.04	13.73	**0.14**	**18.31**
Affect and Wellbeing												
Pos. Affect	0.08	11.95	0.06	6.88	0.08	6.81	0.14	7.22	**0.16**	**7.20**	0.01	49.12
Neg. Affect	**0.09**	**20.94**	-0.05	11.05	0.01	10.81	0.09	11.35	-0.0	11.69	0.02	38.29
Anxiety (STAI)	0.06	16.05	-0.0	9.77	-0.01	9.5	**0.10**	**9.88**	0.15	9.68	0.09	52.22
Sleep Quality (PSQI)	0.05	28.73	0.05	15.53	**0.18**	**11.00**	0.05	16.23	0.05	16.33	0.06	34.74

Table B.5: Physical Sensor based Models: Predicting psychological constructs with only physical sensor based features. Prediction algorithms used include Ridge, Elastic Net (ElNet), Support Vector Regressor (SVR), XGBoost (XGB), Gradient Boosted Random Forest (GBR), and Multilayer Perceptron Regressor (MLP). Reported accuracy numbers are Symmetric Mean Absolute Percentage Error (SMAPE) and Pearson's correlation coefficient (r), which are pooled in k-fold cross-validation (k=5). The bold-faced number in each row indicate the best performing model for that construct.

Construct	Algorithm											
	Ridge		ElNet		SVR		XGB		GBR		MLP	
	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE
Cognitive Ability												
Shipley (Abstraction)	**0.12**	**6.81**	-0.19	6.78	0.03	6.75	0.11	7.17	0.06	7.31	-0.06	8.01
Shipley (Vocabulary)	0.10	4.33	-0.14	4.33	**0.12**	**4.13**	-0.0	4.76	0.03	4.77	-0.02	5.62
Personality Traits												
Openness	0.00	6.87	-0.13	6.65	-0.05	6.74	**0.02**	**7.02**	0.01	7.01	-0.02	8.05
Conscientiousness	0.12	7.89	-0.14	7.07	**0.13**	**7.86**	0.12	7.29	0.12	7.21	-0.01	8.21
Extraversion	**0.17**	**8.41**	-0.14	8.69	0.13	8.43	0.12	8.69	0.16	8.69	0.12	9.3
Agreeableness	0.08	5.94	-0.15	5.78	**0.10**	**5.73**	0.03	6.09	0.0	6.12	-0.03	7.13
Neuroticism	**0.12**	**12.94**	-0.17	13.17	0.13	12.93	0.13	13.13	0.12	13.22	0.08	14.34
Affect and Wellbeing												
Pos. Affect	**0.11**	**7.90**	-0.07	6.88	0.11	7.78	0.14	7.03	0.10	7.06	0.08	7.10
Neg. Affect	**0.09**	**11.88**	-0.17	10.89	-0.01	10.81	0.06	11.42	0.08	11.47	0.08	11.26
Anxiety (STAI)	0.11	9.45	-0.14	9.66	0.04	9.50	0.09	9.98	**0.14**	**9.78**	0.09	9.92
Sleep Quality (PSQI)	**0.17**	**16.28**	-0.14	15.52	0.13	15.96	0.15	16.63	0.14	16.71	0.06	16.94

Table B.6: Mixed-effects Models: Predicting psychological constructs with both physical activity and social media features together. Prediction algorithms used include Ridge, Elastic Net (ElNet), Support Vector Regressor (SVR), XGBoost (XGB), Gradient Boosted Random Forest (GBR), and Multilayer Perceptron Regressor (MLP). Reported accuracy numbers are Symmetric Mean Absolute Percentage Error (SMAPE) and Pearson's correlation coefficient (r), which are pooled in k-fold cross-validation (k=5). The bold-faced number in each row indicate the best performing model for that construct.

Construct	**Algorithm**											
	Ridge		**ElNet**		**SVR**		**XGB**		**GBR**		**MLP**	
	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE	r	SMAPE
Cognitive Ability												
Shipley (Abstraction)	0.15	11.96	-0.19	6.78	**0.25**	**6.65**	0.16	6.96	0.18	6.95	-0.02	22.2
Shipley (Vocabulary)	0.14	8.01	-0.11	4.33	0.25	4.10	0.26	4.36	**0.28**	**4.33**	0.12	29.37
Personality Traits												
Openness	0.13	11.14	-0.13	6.65	**0.26**	**6.40**	0.16	6.65	0.17	6.63	0.16	15.13
Conscientiousness	0.09	10.93	-0.14	7.07	**0.17**	**7.86**	0.17	8.03	0.17	7.99	0.07	15.15
Extraversion	0.03	13.98	-0.14	8.69	0.17	8.37	**0.17**	**8.35**	0.17	8.57	0.09	17.52
Agreeableness	0.07	9.88	-0.15	5.78	**0.17**	**5.91**	0.16	5.92	0.11	5.94	0.14	14.37
Neuroticism	0.10	21.18	-0.17	13.17	**0.11**	**12.93**	0.07	13.45	0.08	13.49	0.13	19.59
Affect and Wellbeing												
Pos. Affect	0.07	11.19	-0.0	6.87	**0.13**	**6.77**	0.13	6.95	0.13	6.85	0.07	29.51
Neg. Affect	0.13	17.82	-0.17	10.89	-0.05	10.83	**0.13**	**11.18**	0.11	11.23	0.12	22.95
Anxiety (STAI)	**0.13**	**16.07**	-0.08	9.65	0.02	9.50	-0.0	10.15	0.07	9.91	0.13	30.71
Sleep Quality (PSQI)	0.13	27.64	-0.14	15.52	0.20	14.9	0.19	15.15	0.19	15.28	**0.21**	**23.22**

APPENDIX C

LIFE EVENTS DISCLOSURES ON SOCIAL MEDIA

C.1 Additional regression models

Disentangling Factors of Reporting Life Events on Different Modalities

Besides the convergence ($\mathbf{Model_1}$) and divergence models ($\mathbf{Model_2}$) as studied in Section subsubsection 6.2.3, we also run a third kind of logistic regression models on the entire data of $\mathtt{D_T}$, such that:

- $\mathbf{Model_{3a}}$ uses all the described covariates as dependent variable and predicts if the event is disclosed on social media as the dependent variable, i.e., 1 if self-disclosed on social media, and 0 if not.

- $\mathbf{Model_{3b}}$ uses all the described covariates as dependent variable and predicts if the event is reported on survey as the dependent variable, i.e., 1 if reported on survey, and 0 if not.

Essentially, these models allow us to disentangle the effects of each of our covariates in explaining the direction of reporting, treating each of the modalities independent of each other. For instance, $\mathtt{Model_2}$ revealed that males show a negative correlation (Table 7) which could either be because males tend to disclose lesser on social media, or because Males report more on surveys compared to females. The two models $\mathbf{Model_{3a}}$ and $\mathbf{Model_{3b}}$ would help us to disentangle similar directions of the factors in each of the models.

Table C.1 shows standardized coefficients and significance of the covariates in the above models. Looking at the significant variables, we find that an interesting pattern that $\mathbf{Model_{3a}}$ and $\mathbf{Model_{3b}}$ show coefficients with opposite signs. For example, age shows positive association with social media disclosures and a negative association with survey

Table C.1: $\mathbf{Model_{3*}}$: Coefficients of linear regression of relevant covariates as independent variables and disclosing on social media as dependent variable in $\mathbf{Model_{3a}}$ (1 for disclosure and 0 for no-disclosure), and self-reporting on survey as dependent variable in $\mathbf{Model_{3b}}$ (1 for self-report and 0 for no-self-report), * p<0.05, ** p<0.01, *** p<0.001.

	$\mathbf{Model_{3a}}$		$\mathbf{Model_{3b}}$			$\mathbf{Model_{3a}}$		$\mathbf{Model_{3b}}$	
Demographic/Trait	**Coeff.**	**p**	**Coeff.**	**p**	**Event Attribute**	**Coeff.**	**p**	**Coeff.**	**p**
Age	0.02	*	−0.04	***	Valence: Positive	0.39	***	−0.19	
Gender: Male	−0.89	*	0.43	***	Significance	−1.26	***	0.87	***
Born in US: Yes	−0.35		0.22		Recency	−1.88		1.57	***
Education: H. School	−0.02		1.16	***	Anticipated	0.19	*	−0.16	
Education: College	−0.04		1.19	*	Intimacy	−0.78	***	0.36	***
Education: Grad School	0.22		0.94	**	Scope	−0.84	***	0.44	**
Education: Doctoral	0.32		0.72		Status: Ongoing	4.71	***	−1.92	***
Shipley: Abstraction	−0.07	***	0.04	*	Type: Health	−0.25		0.29	
Shipley: Vocabulary	−0.02		0.05	**	Type: Work	−1.54	***	0.97	***
Personality: Openness	0.06		−0.34	**	Type: School	−0.10		0.26	
Personality: Conscientiousness	−0.06		−0.07	*	Type: Local	−1.08	*	−0.17	
Personality: Extraversion	0.22	*	−0.03		Type: Financial	−2.88	***	1.35	***
Personality: Agreeableness	0.92	***	0.08						
Personality: Neuroticism	0.26	*	−0.03		**Baseline Attribute**	**Coeff.**	**p**	**Coeff.**	**p**
Positive Affect	−0.00		−0.03	**	SM: Num. Posts	1.28	***	-0.03	
Negative Affect	−0.00		0.00		SM: Avg. Post Length	8.62	*	-1.44	
Stai: Anxiety	0.02		−0.03	*	SR: Num. Records	-1.07	***	1.46	***
PSQI: Healthy Sleep Quality	−0.10	***	0.01		SR: Avg. Significance	0.50	***	-0.17	**

$\mathbf{Model_{3a}}$: AIC $= 920.3$, Deg. Freedom$=34$, LLk. $= -425.15$, $\chi^2 = 2493.88$, Pseudo $R^2 = 0.75$, $p < 0.001$ ***
$\mathbf{Model_{3b}}$: AIC $= 2160.4$, Deg. Freedom$=34$, LLk. $= -1045.22$, $\chi^2 = 1468.90$, Pseudo $R^2 = 0.43$, $p < 0.001$ ***

self-reports. Again, males are less likely to disclose events on social media, and, age has no effect on self-reports. We also find that healthy sleep quality has a strong negative association with social media disclosures, however no significant association with self-reports of life events.

Among event attributes, we find that valence of event bears a strong positive association with social media disclosures but no significant relationship with self-reports. In contrast, greater the significance of an event, less likely it is to be disclosed on social media, and more likely it is to be reported in self-reported survey. We construe similar explanation as in Section subsection 6.2.3 holds here, significant events could be associated with emergency circumstances when the individual has lower propensity to post about the event. Similar associations are observed for recency, intimacy, and scope, with negative association with social media disclosure and positive association with self-reports. With respect to type of events, Work shows significant negative relationship with social media disclosure and positive relationship with self-reports — indicating that work related events are less likely to be posted on social media despite their occurrences.

Finally, we also find interesting directions for the baseline attributes, we find that social media related baseline attributes positively associate with social media disclosure but show no statistical significance in the relationship with survey based disclosure. For survey related baseline attributes, we find that number of survey records negatively associate with number of social media disclosures, and positively associate with survey event logging. Again, baseline self-reported significance shows a positive association with social media disclosure, indicating that individuals who tend to self-perceive greater significance of events are also more likely to disclose the event on social media. Taken together, the relationships observed in this analysis is not very different from what we observe in our results, providing more insight about what does the factors associated with online disclosures of life events.

C.2 Codebooks

Table C.2: Codebook describing definitions and characteristics of identifying life events on social media

Category	Definition
School	
Back To School	• Either the poster or an immediate family member started at a new school (after graduation) or training program (general categories of preschool, K-12, college, etc). • Immediate family members include people that the poster lives with such as partner or children
Changed School	• The poster or immediate family member changed schools or training programs. • Includes change from elementary to middle school and middle to high school, such as finished middle school and will now be going to high school.
Finished School	• Either the poster or their immediate family member graduated elementary/middle/high school/university • Milestone related to graduation (senior prom, baccalaureate, etc). • The poster or their immediate family member graduated/left an organization important to them that they were part of for long time (e.g. a sport) • Finished daycare/preschool and will now be going to K-12 • Completed training/got certified for something (e.g., scuba diving)
Issue at School	• Either the poster or an immediate family member had problems at their school/training program • Very poor grades • In trouble with teacher/principal for something serious • Another person at their school (or family member's school) did something that impacted the poster (this does not include school shooting— that goes under "Assaulted" category).
Failed School	• The poster or their immediate family member failed school or a training program • Failure could be either completely failing out of school/the program or a significant failure in a class
Did not Finish School	• The poster or their immediate family member did not graduate
Personal	
Engaged	• Got engaged • Talking about plans for marriage, but does not seem like their wedding has been planned yet (no date set, wedding isn't for a long time, etc)
Broken Engagement	• They ended their engagement.
Married	• Posted about wedding Post about their upcoming wedding (have a set date, wedding preparations like buying wedding dress, tasting cake, etc) • Posting about their recent wedding • Could also be an immediate family member's wedding life event of gaining a new in-law, stepparent, etc.
Started Affair	• Started an extramarital affair.

Negative Relationship	• Having serious problems in their relationship (but not separating or getting a divorce)
Separated In Marriage	• They are separating from their partner. • Ended a long-term, significant relationship • Their parents are separating
Divorce	• They are getting divorced. • Their parents are getting divorced.
Positive Relationship	• Partner made a significant (romantic) gesture. • Post about anniversary or other important, positive relationship milestone. (For anniversary, people may use expressions like "another year with") • Couple participated in hobby, activity, etc together • Partner being supportive of them during a difficult time.
Reunion After Separation	• They get back with their partner after being separated or divorced.
Infidelity	• Either they cheated on their partner or their partner cheated on them.
Trouble In-Laws	• Posted about difficulties with their in-laws
Spouse Died	• Posted about spouse's death or funeral
Pregnant	• Announced their or their partner's pregnancy. • They posted something significant related to their current pregnancy (e.g. first time baby kicked). • Do not include things that inconclusive events (e.g. morning sickness, having weird cravings) • Do not include posts reminiscing about past pregnancy/pregnancies
FirstChild	• Announcing the birth of first child. • Assume it is first child if no other children mentioned. • Use for when infant has significant milestones important to parent within first year— crawling, walking. • Can use up until the child turns one-year old (1st birthday can be included) • Becomes a grandparent"
Younger Child Birth	• They had another child. • Use for when infant has significant milestones important to parent within first year— crawling, walking. • Can use up until the child turns one-year old (1st birthday can be included)
FertilityIssue	• They learned that they are not able to have children. • Had surgery to prevent having children.
Child Died	• Their child died (may mention "SIDS" - a common cause) • Do NOT include abortions or miscarriages here— they have their own categories in the health-related section. • Can still categorize with this even if a lot of time has passed since it's an event that continues to cause the person great pain.
Adopted Child	• Adopted child(ren) • Fostering child(ren) • Became a stepmother or stepfather
Person Move In	• Someone moved into their household.

Person Moved Out	• Someone moves out of the poster's home. • If post is about child going off to college- chose this instead of "Back To School" (because this is more directly relevant for the poster).
Person Stayed Longer	• Someone stayed in their home after they were supposed to move out.
Argument In Family	• Serious argument with a family member that is not partner or spouse.
More/Less Family Meetups	• **More:** • See family member(s) after a long time apart • Meet a family member's new baby (sister's baby, cousin's baby. . .) • A family member gets married — the life event of gaining a new in-law, stepparent, etc. • **Less:** • Have not seen one or more family members in a long time"
Death in Family	• Family member died (recently enough that grief is still fresh) • Family member with dementia— feels like they're already gone.
Positive Move	• Happy about recent move to a new residence or neighborhood (if there are exclamation points and the tone is positive, can assume Positive Move) • About to move to a new residence or neighborhood • Moved to a better residence/neighborhood than old one. • Become a citizen or earn residency (e.g., green card in the U.S.)
Negative Move	• Unhappy about their move to a new residence or neighborhood • Having trouble with the moving process.
Neutral Move	• Neutral view about moving (or post is just ambiguous about). • Has "mixed emotions" or calls the move "bittersweet" • Planning to move but have not yet — selling home, made an offer on a house, etc."
Failed Move	• They were unable to move after attempting to. • Complete failure- they are staying in current home, not renting or waiting it out."
Build Home	• Build a new home (or more likely, have it built)
Remodeled Home	• They remodeled at least one room in their home (or a significant feature, like adding a pool).
Lost Home in Disaster	• Part of home was significantly damaged in a disaster. • Lost entire home in disaster.
Assaulted	• Physically assaulted, or involved in a similar traumatizing event. • This includes events that create significant trauma (e.g., in a school during a shooting) • Victim of sexual harassment • Either the poster or an immediate family member.
Robbed	• They were directly robbed, an immediate family member was robbed, or their home was • They were a victim of identity theft, or credit card was stolen, etc.

No Injury Accident	• Involved in an accident but not injured (or barely injured— just a few bruises, etc) • Witness to a big accident (e.g. a deadly car crash) • Helps out someone else who was involved in a serious accident • Very worried about a loved one/pet (e.g they are missing) • Hurricane, wildfire, etc impacting their city
Lawsuit	• They become involved in a lawsuit.
Accused	• They are accused of a crime.
Lost License	• They lost their driver's license • They lose some other form of license (eg. medical license)
Arrested	• They were arrested
Went Jail	• They went to jail/prison. • An immediate family member went to jail/prison. • Someone who harmed them went to jail/prison.
Court Case	• They became involved in a court case. • They were a witness for a court case. • They had jury duty (do not include if they were only summoned- must have at least reported for duty).
Convicted	• They were convicted of a crime. • An immediate family member was convicted. • A person that harmed them was convicted.
Acquitted	• They were acquitted. • An immediate family member was acquitted. • A person that harmed them was acquitted.
Released Jail	• They were released from jail/prison. • An immediate family member was released from jail/prison. • Someone who harmed them was released from jail/prison
In Jail Longer	• Didn't get out of jail/prison when expected (i.e. was up for parole). • An immediate family member didn't get out of jail/prison when expected. • Someone who harmed them did not get released from jail/prison when expected.
Increased Social Activity	• Increase in organizational/hobby-related activity. • Took part in an important event related to this organization, hobby, etc • Or significant increase in time spent with friend(s) (e.g. "girls weekend") • Important reunion (e.g. high school) • Went to a special event • A good friend has an important event — gets married, has a baby, graduates. . . • Running a marathon, big sports event, etc.

Vacation	• Had a holiday away from home (not related to work) • Generally more than a day. Usually if beach is mentioned (unless something in post indicates otherwise) • May still be labeled as “vacation” even if person calls it a “trip,” as long as it meets the criteria • Posting about a popular tourist location— especially if wording indicates they are a tourist. • Cruise • Mentions upcoming vacation, but with some sort of planning
Vacation Plan Fail	• Unexpected events ruining either part or all of vacation plans, e.g., flight cancellation, missed flight, weather, family issue, etc. • Do NOT include layovers unless they are an unexpected, long delay.
New Hobby	• Picked up a new skill, hobby, or craft • New recreational activity • Healthier lifestyle (aka incorporated new healthy habits) has had a significant impact (e.g. new diet, exercise regimen has led to weight loss, feeling better, etc) • Got a tattoo They or their child got driver’s license
Dropped Hobby	• They dropped a hobby that was significant to them (e.g. stopped running due to an injury) • Got rid of an unhealthy habit
New Pet	• Adopted a new pet • Began fostering new pet • Began training a service dog • Instead of directly mentioning “a pet”, one may indicate by including emoji of the type of pet or call an explicit nickname or attributes indicating a pet (e.g., “kitty”, “doggy”, “paws”).
Pet Died	• Their pet died. • They were forced to give away their pet for some reason. • Their pet is very sick/frail and they think it is going to pass away soon.
New Friends	• "Made new friend(s) • Met a celebrity, figure, etc who’s important to them • Share an impactful moment/event with a stranger
Broken Friendship	• Broke up with a close friend In a huge fight with a close friend
Breakup	• Broke up with significant othe • Or planning (or strongly considering) to break up with significant other. • Post is centered around the break up— not just describing someone as an ex when talking about another event (see example of post that does not belong). • Does not include divorce or friendship posts.
Friend Died	• A close friend died • If a death significant to person is mentioned and not sure of relationship, this is the default option.

Trip	• Time away from home that is not for pleasure— likely work related. • Or a quick day trip away from hometown. • Assume "Vacation" if cannot confirm "Trip" • Camping/hiking/day trip. • A sports trip they are participating in. • Posting about the flight they're on.
Work-Related	
First Job	• Started first job. Applies to poster/immediate family
Back to Work	• They returned to work after a long period of not working. • They returned to work after being a stay-at-home parent for a long time. • Not the first job • Applies to poster/immediate family • Assume this when the post doesn't specify if it is a new job
Positive Job Switch	• Happy about leaving job for another. Leaving job to become stay-at-home parent (considered a job).
Negative Job Switch	• Upset or mad about leaving current job for another. • Upset about a recent job switch
Neutral Job Switch	• Neutral feelings about about leaving current job for another.
Boss Trouble	• Had (or having) problems with a boss. • Their boss did something bad in general and/or was fired.
Demoted	• Demoted at work
No Promotion	• Did not get a promotion at work
Bad Work Life	• Significant problems at work (e.g. being harassed by a coworker) • Post about why they really dislike their current job • Don't include minor things like a singular bad day. • Consistently frustrated with job, coworkers.
Promoted	• Got promoted at work.
Work Success	• Significant success at work (not including promotion) • Takes on a leadership role in an event (e.g. speak at a conference, hold an event) • Received an award at a competition. • They got a big bonus. • They had a big commission (if works in sales). • Takes on a leadership role in an event, or represents their company at an event (e.g. speak at a conference) • Hold an event (not necessarily for company— e.g. host a fundraiser for a charity) • Reach a big milestone at work (e.g. 10 year anniversary of working at company)
Good Work-life	• Positive experience(s) at work (significant to them or to the company as a whole). • Conditions improved at work.

Laid Off	• Let go from their job (due to company reasons, like financial problems)
Fired	• Fired from their job
Startup	• Started a new business— • Includes small, one-person businesses (e.g. selling own artwork)
Expansion	• Their business is growing, selling more, etc. • Hiring on more people at work.
Heavy Work	• They took on a greatly increased workload. • They have a lot of work/home/other organization work combined
Work Loss	• Made a big mistake at work. Caused their company to lose money • If manager, boss, owner- their company, in general, suffered financial loss, went bankrupt. . . "
Light Work	• Their workload decreased significantly.
Performance Review	• Had a performance review at work— or anticipating one. • Company, product, etc reviewed online, in a magazine.
New Project	• Takes on a new project at work, home, etc. • Something that they will finish (not a hobby)
Retired	• Retired from job
Break from Work	• Leaves job of own volition, and does not mention switch to new job. • Not retiring- likely going back to work eventually
Got Bonus	• Got a bonus at work (expected or unexpected). • Pay raise. • Referring to money only— does not include job promotions (separate category).
In Armed Service	• They entered the armed services (Army, Navy, Air Force. . .)
Out Armed Service	• They left the armed services (Army, Navy, Air Force. . .) • Honorable/Dishonorable discharge?
Financial-Related	
Mortgage	• Took out a mortgage • Use if they specifically say that they bought a house
Installment Purchase	• Began or finished paying for a large purchase on an installment plan (e.g. car) • Made a significant purchase (e.g. car, etc)
Mortgage Closed	• Foreclosure of mortgage • Foreclosure of loan
Rebought On Installment	• Repossession of large purchase bought on installment plan (car, furniture. . .)
Salary Cut	• Had a cut in salary/wages, but were not demoted.
Financial Loss	• Suffer a significant personal (non-work related) financial loss. • Categorize as “Financial Loss” if about a wedding, trip, etc if the post is focused on the big financial loss caused by this event.
On Welfare	• Went on welfare
Off Welfare	• Went off welfare
Salary Increment	• Got a substantial increase in wage or salary without a promotion (work related?)

No Salary Increment	• Did not get an expected wage or salary increase (work related?)
Non-work Financial Gain	• Had a significant non-work related financial gain. • Finished paying off a large debt (e.g. student debt, credit card. . .) • Renting out their home, guest house, a room in their home, etc
Health-Related	
Abortion	• Poster or their partner had an abortion (or is planning to)
Miscarriage / Stillbirth	• Had a miscarriage • Had a stillbirth • Does not matter how long ago— still a significant enough event to keep affecting and remembering about it in the present.
Menopause	• They began menopause. • Unlikely to explicitly state this, may be a vague post (talking about being unable to have kids anymore, hot flashes, etc).
Health Gain	• Physical health improvement (from illness, injury, etc.) • Made significant, lasting healthy changes, eating better, exercising, etc. • Mental health improvement (benefited from therapy, new meds working well, etc) • If immediate family member (or person they are very close to) has a significant gain that affects poster, then include (e.g. finally leaving hospital after watching over bedside).
Health Loss	• Decline in physical health (diagnosed with disease, etc— NOT including a common cold or virus) • Big, overall health problem. • Or a decline in mental health (depression, anxiety, etc) • If immediate family member (or person they are very close to) has a very serious problem that also affects the poster (e.g. cancer) include them."
Injury	• Significant physical injury (not just a bruise or cut) • Pain specific to a certain area (or areas) of the body. • Problems with legs, knees, back, etc. • Painful symptoms • Pain related to having surgery • If immediate family member (or person they are very close to) has a very serious injury that affects poster, then include them.
No Treatment	• Unable to get treatment for an illness or injury because the poster 1) does not have insurance, 2) finds the medication is too expensive, 3) does not have the time to see a doctor, etc.
Loccal	
Bad Weather	• Post is complaining about extreme temperatures • Worried about bad weather (rainstorm tornado, snow, wildfire, etc) in their area. • Post focused on how they were negatively affected by a bad weather event (e.g. couldn't drive because of snow).

Sports Event	• Participating in a sports event (marathon, soccer, basketball) • Training for a sports event (e.g. marathon) • Coaching a sports team. • Attending a sports game. • Party/event centered around a sporting event (e.g. Super Bowl Party) • Does not include going to the gym (categorized as hobby) • Not a sporting event if there’s no physical activity involved (e.g. trivia)

Table C.3: Paraphrased example continuous life events identified in our data.

Example posts
Series of posts together
Survey: has anybody had the last six weeks of their pregnancy feel like a constant hunger fest?
Hospital swag? #4weeksout
To all my nurses....
The baby made me eat it.....
Future song writer? Might need some lesson... #likebabylikemomma
Sometimes the thing I get the most excited about is sweatpants. All day long... #preggolife #nothingnfitsright
Glucose testing day! 11 weeks until brother!
Surrounding posts providing context
Hilo, HI on the big island. Amazing landscape!!
Miguel in Hawaii
Room with a view, the views were magnificent!
Zip lining, so fun!!
Talk about fun!! Something I thought I would never do, now I may have to get certified!!
Single post describing a continuous life event
John came home from the hospital last night. Thank you so much for your prayers, love and support. He has 6 weeks of IV antibiotics but at least he gets to do that at home...

Table C.4: A codebook to label intimacy of life events on a Three-point Likert scale.

Degree of Intimacy	Description
Low	Events which are typically discloseable to a public audience, smaller events as a part of normal lives, posts which are common to be post about on social media, e.g., Work and academic success, social activities, family meetups, celebratory post about family, friends, significant other, event that is not too specific, and several others may be undergoing at the same time (e.g., back to school, change of quarter related events), promoting business
Medium	Big life event that is typically shared on social media. (e.g. new baby, new job). Includes negative events that are common to post about (death of family member, injury). Large purchases such as a house, ca. Important part of personal life, but not too personal to share (would be fine telling someone they just met about)
High	Traumatic or events associated with some form of stigma. Sharing negative parts of life. Disclosing events about which someone might feel embarrassed about. Posts that may not make one look "perfect" if disclosed publicly. Finances— considered unusual (sometimes rude) to talk about money (at least in America)

Table C.5: A codebook to label scope (degree of directness on the individual) of life events on a Three-point Likert scale.

Degree of Scope	Description
Low	• They are removed from the event, e.g., describing something in surroundings (e.g. weather event, how nice their workplace is). • Promotional— want to reach as many people as possible. • Big event, not intimate/will interact with strangers, e.g., sports event. • Unlikely to explicitly state this, may be a vague post (talking about being unable to have kids anymore, hot flashes, etc).
Medium	• Both the poster and family, friends, and/or significant other are involved in/were impacted by the event. • Event is happening to someone close to them, but the poster is not part of that event, e.g., School-related posts typically parents talking about children. • Events such as buying a house, moving, etc which also involve their family. • Post about pets often mention family (or consider pet as family)
High	• Event is unique to the poster—– it only happened to them • Or, significantly impacted them more than others, • Examples include promotion, met with an accident themselves, diagnosis about some condition of their own.

Table C.6: A codebook to label temporal status of life events on binary values of ongoing and ended.

Status	Description
Ongoing	• Events that are still going on, or the individual talks about something that is happening in the present. • Part of an continuous or long-term process (e.g. pregnancy-related, planning upcoming wedding, health problems) • May use past tense— but still ongoing if the overall event has not ended. • If they're using past tense, but posting about an event that just happened.
Ended	• Events that recently ended, or ended in the past and the individual is mentioning about the occurrence. • The individual reflects on an event that happened a while ago, but still affects them (e.g. death of a parent). • The part of a continuous event which just ended (e.g. got back from vacation recently).

REFERENCES

[1] A. Abadie, "Semiparametric difference-in-differences estimators", *Rev. Econ. Stud.*, vol. 72, no. 1, pp. 1–19, 2005.

[2] M.-A. Abbasi, S.-K. Chai, H. Liu, and K. Sagoo, "Real-world behavior analysis through a social media lens", in *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, Springer, 2012, pp. 18–26.

[3] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context and context-awareness", in *International Symposium on Handheld and Ubiquitous Computing*, Springer, 1999, pp. 304–307.

[4] D. J. Abramis and L. Beach, "Work Role Ambiguity , Job Satisfaction , and Job Performance : Meta-Analyses and Review", no. February, 2017.

[5] M. S. Ackerman, "The intellectual challenge of cscw: The gap between social requirements and technical feasibility", *Human–Computer Interaction*, vol. 15, no. 2-3, pp. 179–203, 2000.

[6] A. Acquisti and R. Gross, "Imagined communities: Awareness, information sharing, and privacy on the facebook", in *International workshop on privacy enhancing technologies*, Springer, 2006, pp. 36–58.

[7] P. Adams, M. Rabbi, T. Rahman, M. Matthews, A. Voida, G. Gay, T. Choudhury, and S. Voida, "Towards personal stress informatics: comparing minimally invasive techniques for measuring daily stress in the wild", in *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014, pp. 72–79.

[8] I. Ajzen, "From intentions to actions: A theory of planned behavior", in *Action control*, Springer, 1985, pp. 11–39.

[9] I. Ajzen *et al.*, "The theory of planned behavior", *Organizational behavior and human decision processes*, vol. 50, no. 2, pp. 179–211, 1991.

[10] F. Al Zamal, W. Liu, and D. Ruths, "Homophily and latent attribute inference: Inferring latent attributes of twitter users from neighbors", in *Sixth International AAAI Conference on Weblogs and Social Media*, 2012.

[11] N. Albalooshi, N. Mavridis, and N. Al-Qirim, "A survey on social networks and organization development", in *2012 International Conference on Collaboration Technologies and Systems (CTS)*, IEEE, 2012, pp. 539–545.

[12] P. Aldhous and C. Dweck, *Free your mind and watch it grow*, 2008.

[13] C. N. Alexander Jr and P. Lauderdale, “Situated identities and social influence”, *Sociometry*, pp. 225–233, 1977.

[14] T. Althoff, P. Jindal, and J. Leskovec, “Online actions with offline impact: How online social networks influence online and offline user behavior”, in *Proceedings of the tenth ACM international conference on web search and data mining*, 2017, pp. 537–546.

[15] I. Altman and D. A. Taylor, *Social penetration: The development of interpersonal relationships.* Holt, Rinehart & Winston, 1973.

[16] M. Alvesson and S. Sveningsson, *Changing organizational culture: Cultural change work in progress*. Routledge, 2015.

[17] A. Anagnostopoulos, R. Kumar, and M. Mahdian, “Influence and correlation in social networks”, in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 2008, pp. 7–15.

[18] N. Andalibi, “Disclosure, privacy, and stigma on social media: Examining non-disclosure of distressing experiences”, *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 27, no. 3, pp. 1–43, 2020.

[19] N. Andalibi and A. Forte, “Announcing pregnancy loss on facebook: A decision-making framework for stigmatized disclosures on identified social network sites”, in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–14.

[20] N. Andalibi, O. L. Haimson, M. De Choudhury, and A. Forte, “Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity”, in *Proceedings of the 2016 CHI conference on human factors in computing systems*, 2016, pp. 3906–3918.

[21] I. Andone, K. Błaszkiewicz, M. Eibes, B. Trendafilov, C. Montag, and A. Markowetz, “How age and gender affect smartphone usage”, in *Proc. ACM international joint conference on pervasive and ubiquitous computing: adjunct*, 2016.

[22] J. Angst and G. Cassano, “The mood spectrum: improving the diagnosis of bipolar disorder”, *Bipolar disorders*, vol. 7, no. s4, pp. 4–12, 2005.

[23] B. Al-Ani, G. Mark, and B. Semaan, “Blogging in a region of conflict: Supporting transition to recovery”, in *Proceedings of the SIGCHI Conference on human factors in computing systems*, 2010, pp. 1069–1078.

[24] M. M. Antony, P. J. Bieling, B. J. Cox, M. W. Enns, and R. P. Swinson, "Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample.", *Psychological assessment*, vol. 10, no. 2, p. 176, 1998.

[25] A. Archambault and J. Grudin, "A longitudinal study of facebook, linkedin, and twitter use", in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2012, pp. 2741–2750.

[26] R. M. Arkin and J. A. Shepperd, "Strategic self-presentation: An overview.", 1990.

[27] S. Atkinson, A.-M. Bagnall, R. Corcoran, J. South, and S. Curtis, "Being well together: Individual subjective and community wellbeing", *Journal of Happiness Studies*, vol. 21, no. 5, pp. 1903–1921, 2020.

[28] J. W. Ayers, B. M. Althouse, E. C. Leas, T. Alcorn, and M. Dredze, "Can Big Media Data Revolutionarize Gun Violence Prevention?", *arXiv preprint arXiv:1611.01148*, 2016.

[29] A. D. Baddeley and G. Hitch, "The recency effect: Implicit learning with explicit retrieval?", *Memory & Cognition*, vol. 21, no. 2, pp. 146–155, 1993.

[30] G. Bagheri, R. Zarei, and M. N. Aeen, "Organizational silence (basic concepts and its development factors)", *Ideal Type of Management*, vol. 1, no. 1, pp. 47–58, 2012.

[31] S. Bagroy, P. Kumaraguru, and M. De Choudhury, "A Social Media Based Index of Mental Well-Being in College Campuses", in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver, Colorado, USA, 2017.

[32] A. B. Bakker, E. Demerouti, and L. Lieke, "Work engagement, performance, and active learning: The role of conscientiousness", *Journal of vocational behavior*, vol. 80, no. 2, pp. 555–564, 2012.

[33] A. Baldauf, H. Reisinger, and W. C. Moncrief, "Examining motivations to refuse in industrial mail surveys", *Journal of the Market Research Society*, vol. 41, no. 3, pp. 345–354, 1999.

[34] T. Baldwin, P. Cook, M. Lui, A. MacKinlay, and L. Wang, "How noisy social media text, how diffrnt social media sources?", in *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, 2013, pp. 356–364.

[35] A. Bandura, "Organisational applications of social cognitive theory", *Australian Journal of management*, vol. 13, no. 2, pp. 275–302, 1988.

[36] ——, "Perceived self-efficacy in cognitive development and functioning", *Educational psychologist*, vol. 28, no. 2, pp. 117–148, 1993.

[37] ——, "Social cognitive theory of personality", *Handbook of personality*, vol. 2, pp. 154–196, 1999.

[38] ——, "Social cognitive theory in cultural context", *Applied psychology*, vol. 51, no. 2, pp. 269–290, 2002.

[39] S. R. Barley, G. W. Meyer, and D. C. Gash, "Cultures of culture: Academics, practitioners and the pragmatics of normative control", *Administrative science quarterly*, pp. 24–60, 1988.

[40] S. B. Barnes, "A privacy paradox: Social networking in the United States", *First Monday*, vol. 11, no. 9, 2006.

[41] J. N. Baron and W. T. Bielby, "The proliferation of job titles in organizations", *Administrative Science Quarterly*, pp. 561–586, 1986.

[42] M. R. Barrick and M. K. Mount, "The big five personality dimensions and job performance: A meta-analysis", *Personnel psychology*, vol. 44, no. 1, pp. 1–26, 1991.

[43] M. R. Barrick, L. Parks, and M. K. Mount, "Self-monitoring as a moderator of the relationships between personality traits and performance", *Personnel psychology*, vol. 58, no. 3, pp. 745–767, 2005.

[44] M. K. Bartholomew, S. J. Schoppe-Sullivan, M. Glassman, C. M. Kamp Dush, and J. M. Sullivan, "New parents' facebook use at the transition to parenthood", *Family relations*, vol. 61, no. 3, pp. 455–469, 2012.

[45] Y. Baruch and B. C. Holtom, "Survey response rate levels and trends in organizational research", *Human relations*, vol. 61, no. 8, pp. 1139–1160, 2008.

[46] T. S. Bateman and J. M. Crant, "The proactive component of organizational behavior: A measure and correlates", *Journal of organizational behavior*, vol. 14, no. 2, pp. 103–118, 1993.

[47] V. Battistich and A. Hom, "The relationship between students' sense of their school as a community and their involvement in problem behaviors.", *American journal of public health*, vol. 87, no. 12, pp. 1997–2001, 1997.

[48] S. Bauhoff, "The effect of school district nutrition policies on dietary intake and overweight: a synthetic control approach", *Economics & Human Biology*, vol. 12, pp. 45–55, 2014.

[49] E. P. Baumer, P. Adams, V. D. Khovanskaya, T. C. Liao, M. E. Smith, V. Schwanda Sosik, and K. Williams, “Limiting, leaving, and (re) lapsing: An exploration of facebook non-use practices and experiences”, in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2013, pp. 3257–3266.

[50] N. Bayram and N. Bilgel, “The prevalence and socio-demographic correlations of depression, anxiety and stress among a group of university students”, *Social psychiatry and psychiatric epidemiology*, vol. 43, no. 8, pp. 667–672, 2008.

[51] N. N. Bazarova and Y. H. Choi, “Self-disclosure in social media: Extending the functional approach to disclosure motivations and characteristics on social network sites”, *Journal of Communication*, vol. 64, no. 4, pp. 635–657, 2014.

[52] A. Beaulieu, “Research Note: From co-location to co-presence: Shifts in the use of ethnography for the study of knowledge”, *Social Studies of Science*, vol. 40, no. 3, pp. 453–470, 2010.

[53] M. H. Becker, “The health belief model and sick role behavior”, *Health education monographs*, vol. 2, no. 4, pp. 409–419, 1974.

[54] M. Belkin, D. Hsu, S. Ma, and S. Mandal, “Reconciling modern machine-learning practice and the classical bias–variance trade-off”, *Proceedings of the National Academy of Sciences*, vol. 116, no. 32, pp. 15 849–15 854, 2019.

[55] S. J. Bell and B. Menguc, “The employee-organization relationship, organizational citizenship behaviors, and superior service quality”, *Journal of retailing*, vol. 78, no. 2, pp. 131–146, 2002.

[56] A. Benetos, A. Rudnichi, F. Thomas, M. Safar, and L. Guize, “Influence of heart rate on mortality in a French population: role of age, gender, and blood pressure”, *Hypertension*, vol. 33, no. 1, pp. 44–52, 1999.

[57] M. S. Bernstein, E. Bakshy, M. Burke, and B. Karrer, “Quantifying the invisible audience in social networks”, in *Proceedings of the SIGCHI conference on human factors in computing systems*, ACM, 2013, pp. 21–30.

[58] C. Berrocal, M. A. R. Moreno, M. A. Rando, A. Benvenuti, and G. B. Cassano, “Borderline personality disorder and mood spectrum”, *Psychiatry research*, vol. 159, no. 3, pp. 300–307, 2008.

[59] J. L. Bevan, R. Gomez, and L. Sparks, “Disclosures about important life events on facebook: Relationships with stress and quality of life”, *Computers in Human Behavior*, vol. 39, pp. 246–253, 2014.

[60] S. J. Biddle, K. Fox, and S. Boutcher, *Physical activity and psychological well-being*. Routledge, 2003.

[61] G. Biggio and C. Cortese, "Well-being in the workplace through interaction between individual characteristics and organizational context", *International journal of qualitative studies on health and well-being*, vol. 8, no. 1, p. 19 823, 2013.

[62] M. Bin Morshed, K. Saha, M. De Choudhury, G. D. Abowd, and T. Plötz, "Measuring self-esteem with passive sensing", in *Proceedings of the 14th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2020, pp. 363–366.

[63] M. Bin Morshed, K. Saha, R. Li, S. K. D'Mello, M. De Choudhury, G. D. Abowd, and T. Plötz, "Prediction of Mood Instability with Passive Sensing", *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, 2019.

[64] J. Binder, A. Howes, and A. Sutcliffe, "The problem of conflicting social spheres: Effects of network structure on experienced tension in social network sites", in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2009, pp. 965–974.

[65] M. L. Birnbaum, S. K. Ernala, A. F. Rizvi, M. De Choudhury, and J. M. Kane, "A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals", *Journal of Medical Internet Research*, vol. 19, no. 8, 2017.

[66] C. Blanco, M. Okuda, C. Wright, D. S. Hasin, B. F. Grant, S.-M. Liu, and M. Olfson, "Mental health of college students and their non–college-attending peers: results from the national epidemiologic study on alcohol and related conditions", *Archives of general psychiatry*, vol. 65, no. 12, pp. 1429–1437, 2008.

[67] J. M. Blaxton, C. S. Bergeman, B. R. Whitehead, M. E. Braun, and J. D. Payne, "Relationships among nightly sleep quality, daily stress, and daily affect", *The Journals of Gerontology: Series B*, vol. 72, no. 3, pp. 363–372, 2017.

[68] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation", *JMLR*, vol. 3, no. Jan, pp. 993–1022, 2003.

[69] S. M. Boker, J. L. Rotondo, M. Xu, and K. King, "Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series.", *Psychological methods*, vol. 7, no. 3, p. 338, 2002.

[70] W. C. Borman and S. Motowidlo, “Expanding the criterion domain to include elements of contextual performance”, *Personnel Selection in Organizations; San Francisco: Jossey-Bass*, p. 71, 1993.

[71] S. Boulianne, “Social media use and participation: A meta-analysis of current research”, *Information, communication & society*, vol. 18, no. 5, pp. 524–538, 2015.

[72] C.-C. Bowen, J. K. Swim, and R. R. Jacobs, “Evaluating Gender Biases on Actual Job Performance of Real People: A Meta-Analysis 1”, *Journal of Applied Social Psychology*, vol. 30, no. 10, pp. 2194–2215, 2000.

[73] D. Boyd and K. Crawford, “Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon”, *Information, communication & society*, vol. 15, no. 5, pp. 662–679, 2012.

[74] D. M. Boyd and N. B. Ellison, “Social network sites: Definition, history, and scholarship”, *Journal of computer-mediated Communication*, vol. 13, no. 1, pp. 210–230, 2007.

[75] R. L. Boyd and J. W. Pennebaker, “A way with words: Using language for psychological science in the modern era”, 2015.

[76] J. A. Breaugh and J. P. Colihan, “Measuring facets of job ambiguity: Construct validity evidence.”, *Journal of applied psychology*, vol. 79, no. 2, p. 191, 1994.

[77] S. J. Breckler, “Empirical validation of affect, behavior, and cognition as distinct components of attitude.”, *Journal of personality and social psychology*, vol. 47, no. 6, p. 1191, 1984.

[78] J. W. Brehm, “A theory of psychological reactance.”, 1966.

[79] E. O. Brigham, E. O. Brigham, J. Rey Pastor, R. Pastor, T. M. T. M. Apostol, M. Rodrıéguez, M. R. Rodrıéguez, M. R. Martıénez, C. H. Edwards, D. E. H. Edwards, *et al.*, *The fast Fourier transform and its applications*, 517.443. Prentice Hall, 1988.

[80] E. Brown, *Americans prefer sharing life’s milestones with their social network than in person*, https://www.zdnet.com/article/americans-prefer-sharing-lifes-milestones-with-their-social-network-than-in-person/, Accessed: 2020-09-15, 2018.

[81] T. A. Brown, B. F. Chorpita, and D. H. Barlow, “Structural relationships among dimensions of the DSM-IV anxiety and mood disorders and dimensions of negative affect, positive affect, and autonomic arousal.”, *Journal of abnormal psychology*, vol. 107, no. 2, p. 179, 1998.

[82] J. Brubaker, F. Kivran-Swaine, L. Taber, and G. Hayes, "Grief-stricken in a crowd: The language of bereavement and distress in social media", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 6, 2012.

[83] J. R. Brubaker, G. R. Hayes, and P. Dourish, "Beyond the grave: Facebook as a site for the expansion of death and mourning", *The Information Society*, vol. 29, no. 3, pp. 152–163, 2013.

[84] A. Bryson, J. Forth, and L. Stokes, "Does worker wellbeing affect workplace performance", *Department for Business Innovation & Skills*, 2014.

[85] M. J. Brzozowski, "WaterCooler: exploring an organization through enterprise social media", in *Proceedings of the ACM 2009 international conference on Supporting group work*, ACM, 2009, pp. 219–228.

[86] A. Buchanan and M. L. Kern, "The benefit mindset: The psychology of contribution and everyday leadership", *International Journal of Wellbeing*, vol. 7, no. 1, 2017.

[87] T. Bucher, A. Helmond, *et al.*, "The affordances of social media platforms", *The SAGE handbook of social media*, pp. 233–253, 2017.

[88] A. S. Buchman, R. S. Wilson, L. Yu, B. D. James, P. A. Boyle, and D. A. Bennett, "Total daily activity declines more rapidly with increasing age in older adults", *Archives of gerontology and geriatrics*, vol. 58, no. 1, pp. 74–79, 2014.

[89] C. J. Burges and B. Schölkopf, "Improving the accuracy and speed of support vector machines", *Advances in neural information processing systems*, pp. 375–381, 1997.

[90] M. Burke, J. Cheng, and B. de Gant, "Social comparison and facebook: Feedback, positivity, and opportunities for comparison", in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

[91] M. Burke and R. Kraut, "Using facebook after losing a job: Differential benefits of strong and weak ties", in *Proceedings of the 2013 conference on Computer supported cooperative work*, 2013, pp. 1419–1430.

[92] D. G. Cable, "Grief counseling for survivors of traumatic loss.", 1996.

[93] D. M. Cable and T. A. Judge, "Person–organization fit, job choice decisions, and organizational entry", *Organizational behavior and human decision processes*, vol. 67, no. 3, pp. 294–311, 1996.

[94] C. Cadwalladr and E. Graham-Harrison, "Revealed: 50 million facebook profiles harvested for cambridge analytica in major data breach", *The Guardian*, vol. 17, 2018.

[95] R. Caers and V. Castelyns, "LinkedIn and Facebook in Belgium: The influences and biases of social network sites in recruitment and selection procedures", *Social Science Computer Review*, vol. 29, no. 4, pp. 437–448, 2011.

[96] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, R. A. Peterson, H. Lu, X. Zheng, M. Musolesi, K. Fodor, and G.-S. Ahn, "The rise of people-centric sensing", *IEEE Internet Computing*, vol. 12, no. 4, 2008.

[97] M. A. Campion, E. M. Papper, and G. J. Medsker, "Relations between work team characteristics and effectiveness: A replication and extension", *Personnel psychology*, vol. 49, no. 2, pp. 429–452, 1996.

[98] L. Canzian and M. Musolesi, "Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis", in *Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*, 2015, pp. 1293–1304.

[99] R. D. Caplan and K. W. Jones, "Effects of work load, role ambiguity, and type A personality on anxiety, depression, and heart rate.", *Journal of applied psychology*, vol. 60, no. 6, p. 713, 1975.

[100] J. Caramujo and A. M. R. da Silva, "Analyzing privacy policies based on a privacy-aware profile: The Facebook and LinkedIn case studies", in *2015 IEEE 17th Conference on Business Informatics*, 2015.

[101] C. C. Caruso, T. Bushnell, D. Eggerth, A. Heitmann, B. Kojola, K. Newman, R. R. Rosa, S. L. Sauter, and B. Vila, "Long working hours, safety, and health: toward a National Research Agenda", *American journal of industrial medicine*, vol. 49, no. 11, pp. 930–942, 2006.

[102] R. Catalano, *Health, behavior and the community: An ecological perspective*. Pergamon Press New York, 1979.

[103] R. B. Cattell, "Theory of fluid and crystallized intelligence: A critical experiment.", *Journal of educational psychology*, vol. 54, no. 1, p. 1, 1963.

[104] R. B. Cattell, *Intelligence: Its structure, growth and action*. Elsevier, 1987, vol. 35.

[105] A. Cetto, J. Klier, M. Klier, A. Richter, and K. Wiesneth, "The Blessing of Giving: Kowledge Sharing and Knowledge seeking in Enterprise Social Networks.", in *ECIS*, 2016, ResearchPaper64.

[106] J. Chae, D. Thom, Y. Jang, S. Kim, T. Ertl, and D. S. Ebert, "Public behavior response analysis in disaster events utilizing visual analytics of microblog data", *Computers & Graphics*, vol. 38, pp. 51–60, 2014.

[107] A. Chamberlain, “Does company culture pay off? Analyzing stock performance of’Best Places to Work’companies”, Glassdoor Research Report, March 2015, Tech. Rep., 2015.

[108] ——, “Give to Get: A Mechanism to Reduce Bias in Online Reviews”, Glassdoor Research Report, October 2017, Tech. Rep., 2017.

[109] L. Chan, V. D. Swain, C. Kelley, K. de Barbaro, G. D. Abowd, and L. Wilcox, “Students’ experiences with ecological momentary assessment tools to report on emotional well-being”, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1–20, 2018.

[110] S. Chancellor and S. Counts, “Measuring Employment Demand Using Internet Search Data”, in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM, 2018, p. 122.

[111] S. Chancellor, Z. (Lin, E. Goodman, S. Zerwas, and M. De Choudhury, “Quantifying and Predicting Mental Illness Severity in Online Pro-Eating Disorder Communities”, in *Proceedings of the 19th ACM conference on Computer supported cooperative work & social computing*, in press, ACM, 2016.

[112] S. Chancellor, G. Nitzburg, A. Hu, F. Zampieri, and M. De Choudhury, “Discovering alternative treatments for opioid use recovery using social media”, in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–15.

[113] E. Chang and K. Hancock, “Role stress and role ambiguity in new nursing graduates in Australia”, *Nursing & health sciences*, vol. 5, no. 2, pp. 155–163, 2003.

[114] J. Chang, J. Boyd-Graber, C. Wang, S. Gerrish, and D. M. Blei, “Reading tea leaves: How humans interpret topic models”, in *Neural information processing systems*, Citeseer, vol. 22, 2009, pp. 288–296.

[115] O. Chapelle, B. Scholkopf, and A. Zien, “Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]”, *IEEE Transactions on Neural Networks*, vol. 20, no. 3, pp. 542–542, 2009.

[116] C.-K. Chen, Y.-C. Pei, N.-H. Chen, L.-T. Huang, S.-W. Chou, K. P. Wu, P.-C. Ko, A. M. Wong, and C.-K. Wu, “Sedative music facilitates deep sleep in young adults”, *Journal of Alternative and Complementary Medicine*, 2014.

[117] L. F. Chen, C. Carriker, R. Staheli, P. Isaacs, B. Elliott, B. A. Miller, D. J. Anderson, R. W. Moehring, S. Vereen, J. Bringhurst, *et al.*, “Observing and improving hand hygiene compliance: Implementation and refinement of an electronic-assisted direct-observer hand hygiene audit program”, *Infection control and hospital epidemiology:*

the official journal of the Society of Hospital Epidemiologists of America, vol. 34, no. 2, p. 207, 2013.

[118] L. F. Chen, M. W. Vander Weg, D. A. Hofmann, and H. S. Reisinger, "The Hawthorne effect in infection prevention and epidemiology", *Infection control & hospital epidemiology*, vol. 36, no. 12, pp. 1444–1450, 2015.

[119] J. Cheng, M. Burke, and E. G. Davis, "Understanding perceptions of problematic facebook use: When people experience negative life impact and a lack of control", in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–13.

[120] C. M. Cheung, M. K. Lee, and Z. W. Lee, "Understanding the continuance intention of knowledge sharing in online communities of practice through the post-knowledge-sharing evaluation processes", *Journal of the American Society for Information Science and Technology*, vol. 64, no. 7, pp. 1357–1374, 2013.

[121] M. Chiesa and S. Hobbs, "Making sense of social research: How useful is the Hawthorne Effect?", *European Journal of Social Psychology*, vol. 38, no. 1, pp. 67–74, 2008.

[122] E. K. Choe, S. Abdullah, M. Rabbi, E. Thomaz, D. A. Epstein, F. Cordeiro, M. Kay, G. D. Abowd, T. Choudhury, J. Fogarty, *et al.*, "Semi-Automated Tracking: A Balanced Approach for Self-Monitoring Applications", *IEEE Pervasive Computing*, vol. 16, no. 1, pp. 74–84, 2017.

[123] M. Choi and C. L. Toma, "Social sharing through interpersonal media: Patterns and effects on emotional well-being", *Computers in Human Behavior*, vol. 36, pp. 530–541, 2014.

[124] C.-F. Chung, E. Agapie, J. Schroeder, S. Mishra, J. Fogarty, and S. A. Munson, "When personal tracking becomes social: Examining the use of instagram for healthy eating", in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017.

[125] C.-F. Chung, Q. Wang, J. Schroeder, A. Cole, J. Zia, J. Fogarty, and S. A. Munson, "Identifying and planning for individualized change: Patient-provider collaboration using lightweight food diaries in healthy eating and irritable bowel syndrome", *PACM IMWUT*, 2019.

[126] C. Chung and J. W. Pennebaker, "The psychological functions of function words", *Social communication*, pp. 343–359, 2007.

[127] W. S. Cleveland, "LOWESS: A program for smoothing scatterplots by robust locally weighted regression", *The American Statistician*, vol. 35, no. 1, pp. 54–54, 1981.

[128] S. Cobb, *Termination: The consequences of job loss*, 77. US Department of Health, Education, and Welfare, Public Health Service ..., 1988.

[129] J. Cohen, “Statistical power analysis”, *Curr. Dir. Psychol. Sci.*, 1992.

[130] S. Cohen, T. Kamarck, and R. Mermelstein, “A global measure of perceived stress”, *Journal of health and social behavior*, pp. 385–396, 1983.

[131] M. A. Cohn, M. R. Mehl, and J. W. Pennebaker, “Linguistic markers of psychological change surrounding september 11, 2001”, *Psychological science*, vol. 15, no. 10, pp. 687–693, 2004.

[132] B. E. Compas, G. E. Davis, and C. J. Forsythe, “Characteristics of life events during adolescence”, *American Journal of Community Psychology*, vol. 13, no. 6, pp. 677–691, 1985.

[133] R. A. Cooke and D. M. Rousseau, “Behavioral norms and expectations: A quantitative approach to the assessment of organizational culture”, *Group & Organization Studies*, vol. 13, no. 3, pp. 245–273, 1988.

[134] R. A. Cooke and J. L. Szumal, “Measuring normative beliefs and shared behavioral expectations in organizations: The reliability and validity of the Organizational Culture Inventory”, *Psychological reports*, vol. 72, no. 3_suppl, pp. 1299–1330, 1993.

[135] ——, “Using the organizational culture inventory to understand the operating cultures of organizations”, *Handbook of organizational culture and climate*, vol. 4, pp. 1032–1045, 2000.

[136] G. Coppersmith, M. Dredze, and C. Harman, “Quantifying mental health signals in twitter”, in *Proc. ACL CLCP Workshop*, 2014.

[137] G. Coppersmith, M. Dredze, C. Harman, and K. Hollingshead, “From ADHD to SAD: Analyzing the Language of Mental Health on Twitter through Self-Reported Diagnoses”, in *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, Denver, Colorado, USA, June. North American Chapter of the Association for Computational Linguistics*, 2015.

[138] G. Coppersmith, C. Harman, and M. Dredze, “Measuring post traumatic stress disorder in twitter”, in *International AAAI Conference on Web and Social Media*, 2014.

[139] J. M. Cortina and J. N. Luchman, “Personnel selection and employee performance”, *Handbook of Psychology, Second Edition*, vol. 12, 2012.

[140] W. Coryell, W. Scheftner, M. Keller, J. Endicott, J. Maser, and G. L. Klerman, "The enduring psychosocial consequences of mania and depression.", *The American journal of psychiatry*, 1993.

[141] R. W. Cowdry, D. L. Gardner, *et al.*, "Mood variability: A study of four groups", *The American Journal of Psychiatry*, vol. 148, no. 11, p. 1505, 1991.

[142] J. A.-M. Coyle-Shapiro, "A psychological contract perspective on organizational citizenship behavior", *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, vol. 23, no. 8, pp. 927–946, 2002.

[143] J. M. Crant, "Proactive behavior in organizations", *Journal of management*, vol. 26, no. 3, pp. 435–462, 2000.

[144] M. Crespo and C. E. Lyon, "Employer Access to Employee Social Media: Applicant Screening, "Friend" Requests and Workplace Investigations", *Compliance, Employment Law, Privacy*, 2015.

[145] L. Crocker and J. Algina, *Introduction to classical and modern test theory.* ERIC, 1986.

[146] M. Csikszentmihalyi and R. Larson, "Validity and reliability of the experience-sampling method", in *Flow and the foundations of positive psychology*, Springer, 2014, pp. 35–54.

[147] A. Culotta, "Estimating county health statistics with Twitter", in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2014, pp. 1335–1344.

[148] T. J. Cunningham, C. R. Crowell, S. E. Alger, E. A. Kensinger, M. A. Villano, S. M. Mattingly, and J. D. Payne, "Psychophysiological arousal at encoding leads to reduced reactivity but enhanced emotional memory following sleep", *Neurobiology of Learning and Memory*, 2014.

[149] L. L. Dahlberg and E. G. Krug, "Violence: a global public health problem", *Ciência & Saúde Coletiva*, vol. 11, pp. 1163–1178, 2006.

[150] R. Dara, S. C. Kremer, and D. A. Stacey, "Clustering unlabeled data with SOMs improves classification of labeled real-world data", in *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290)*, IEEE, vol. 3, 2002, pp. 2237–2242.

[151] S. Das and A. Kramer, "Self-censorship on facebook", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 7, 2013.

[152] V. Das Swain, M. D. Reddy, K. A. Nies, L. Tay, M. De Choudhury, and G. D. Abowd, "Birds of a Feather Clock Together: A Study of Person–Organization Fit Through Latent Activity Routines", *Proc. ACM Hum.-Comput. Interact*, no. CSCW, 2019.

[153] V. Das Swain, K. Saha, G. D. Abowd, and M. De Choudhury, "Social media and ubiquitous technologies for remote worker wellbeing and productivity in a post-pandemic world", in *Proceedings of 2020 IEEE 2nd International Conference on Cognitive Machine Intelligence (CogMI)*, 2020.

[154] V. Das Swain, K. Saha, H. Rajvanshy, A. Sirigiri, J. M. Gregg, S. Lin, G. J. Martinez, S. M. Mattingly, S. Mirjafari, R. Mulukutla, S. Nepal, K. Nies, M. D. Reddy, P. Robles-Granda, A. T. Campbell, N. V. Chawla, S. D'Mello, A. K. Dey, K. Jiang, Q. Liu, G. Mark, E. Moskal, A. Striegel, L. Tay, G. D. Abowd, and M. De Choudhury, "A multisensor person-centered approach to understand the role of daily activities in job performance with organizational personas", *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 3, no. 4, 2019.

[155] V. Das Swain, K. Saha, M. D. Reddy, H. Rajvanshy, G. D. Abowd, and M. De Choudhury, "Modeling organizational culture with workplace experiences shared on glassdoor", in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–15.

[156] N. Daskalova, B. Lee, J. Huang, C. Ni, and J. Lundin, "Investigating the effectiveness of cohort-based sleep recommendations", *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 1–19, 2018.

[157] H. Daumé III, "Frustratingly easy domain adaptation", *arXiv preprint arXiv:0907.1815*, 2009.

[158] R. M. Dawes and B. Corrigan, "Linear models in decision making.", *Psychological bulletin*, vol. 81, no. 2, p. 95, 1974.

[159] M. De Choudhury and S. Counts, "Understanding affect in the workplace via social media", in *Proceedings of the 2013 conference on Computer supported cooperative work*, ACM, 2013, pp. 303–316.

[160] M. De Choudhury, S. Counts, and M. Gamon, "Not all moods are created equal! exploring human emotional states in social media", in *Sixth international AAAI conference on weblogs and social media*, 2012.

[161] ——, "Not all moods re created equal! a exploring human emotional states in social media", in *Proc. ICWSM*, 2012.

[162] M. De Choudhury, S. Counts, and E. Horvitz, "Predicting postpartum changes in emotion and behavior via social media", in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2013, pp. 3267–3276.

[163] ——, "Social media as a measurement tool of depression in populations", in *Proceedings of the 5th Annual ACM Web Science Conference*, ACM, 2013, pp. 47–56.

[164] M. De Choudhury, S. Counts, E. Horvitz, and A. Hoff, "Characterizing and predicting postpartum depression from facebook data", in *Proc. CSCW*, 2014.

[165] M. De Choudhury and S. De, "Mental health discourse on reddit: Self-disclosure, social support, and anonymity", in *Eighth international AAAI conference on weblogs and social media*, 2014.

[166] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting depression via social media", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 7, 2013.

[167] M. De Choudhury, E. Kiciman, M. Dredze, G. Coppersmith, and M. Kumar, "Discovering shifts to suicidal ideation from mental health content in social media", in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, 2016, pp. 2098–2110.

[168] M. De Choudhury, A. Monroy-Hernandez, and G. Mark, "“‘narco” emotions: Affect and desensitization in social media during the mexican drug war", in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2014, pp. 3563–3572.

[169] M. De Choudhury, S. S. Sharma, T. Logar, W. Eekhout, and R. C. Nielsen, "Gender and cross-cultural differences in social media disclosures of mental illness", in *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, 2017, pp. 353–369.

[170] J.-E. De Neve, E. Diener, L. Tay, and C. Xuereb, "The objective benefits of subjective well-being", *World happiness report*, 2013.

[171] R. Deshpandé, J. U. Farley, and F. E. Webster Jr, "Corporate culture, customer orientation, and innovativeness in Japanese firms: a quadrad analysis", *Journal of marketing*, vol. 57, no. 1, pp. 23–37, 1993.

[172] T. J. DeStefano, R. N. Mellott, and J. D. Petersen, "A preliminary assessment of the impact of counseling on student adjustment to college", *Journal of College Counseling*, vol. 4, no. 2, pp. 113–121, 2001.

[173] M. A. DeVito, J. Birnholtz, and J. T. Hancock, “Platforms, people, and perception: Using affordances to understand self-presentation on social media”, in *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, 2017, pp. 740–754.

[174] G. Di Stefano, F. Gino, G. P. Pisano, and B. R. Staats, “Making experience count: The role of reflection in individual learning”, *Harvard Business School NOM Unit Working Paper*, no. 14-093, pp. 14–093, 2016.

[175] D. A. Dickey and W. A. Fuller, “Likelihood ratio statistics for autoregressive time series with a unit root”, *Econometrica: Journal of the Econometric Society*, pp. 1057–1072, 1981.

[176] R. A. Dienstbier, “Arousal and physiological toughness: implications for mental and physical health.”, *Psychological review*, vol. 96, no. 1, p. 84, 1989.

[177] J. DiMicco, D. R. Millen, W. Geyer, C. Dugan, B. Brownholtz, and M. Muller, “Motivations for social networking at work”, in *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, ACM, 2008, pp. 711–720.

[178] J. M. DiMicco, W. Geyer, D. R. Millen, C. Dugan, and B. Brownholtz, “People sensemaking and relationship building on an enterprise social network site”, in *2009 42nd Hawaii International Conference on System Sciences*, IEEE, 2009, pp. 1–10.

[179] M. Dimiccoli, J. Marin, and E. Thomaz, “Mitigating bystander privacy concerns in egocentric activity recognition with deep learning and intentional image degradation”, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, pp. 1–18, 2018.

[180] B. S. Dohrenwend and B. P. Dohrenwend, *Stressful life events: Their nature and effects.* John Wiley & Sons, 1974.

[181] B. S. Dohrenwend, A. R. Askenasy, L. Krasnoff, and B. P. Dohrenwend, “Exemplification of a method for scaling life events: The peri life events scale”, *Journal of health and social behavior*, pp. 205–229, 1978.

[182] Y. Doi, M. Minowa, M. Uchiyama, M. Okawa, K. Kim, K. Shibui, and Y. Kamei, “Psychometric assessment of subjective sleep quality using the Japanese version of the Pittsburgh Sleep Quality Index (PSQI-J) in psychiatric disordered and control subjects”, *Psychiatry research*, vol. 97, no. 2-3, pp. 165–172, 2000.

[183] J. S. Donath, “Identity and deception in the virtual community”, *Communities in cyberspace*, vol. 1996, pp. 29–59, 1999.

[184] A. R. T. Donders, G. J. Van Der Heijden, T. Stijnen, and K. G. Moons, “A gentle introduction to imputation of missing values”, *Journal of clinical epidemiology*, vol. 59, no. 10, pp. 1087–1091, 2006.

[185] T. L. Doolen, M. E. Hacker, and E. M. Van Aken, “The impact of organizational context on work team effectiveness: A study of production team”, *IEEE Transactions on Engineering Management*, vol. 50, no. 3, pp. 285–296, 2003.

[186] C. F. Dormann, J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, *et al.*, “Collinearity: a review of methods to deal with it and a simulation study evaluating their performance”, *Ecography*, vol. 36, no. 1, pp. 27–46, 2013.

[187] A. Doryab, J. K. Min, J. Wiese, J. Zimmerman, and J. I. Hong, “Detection of behavior change in people with depression”, AAAI, 2014.

[188] F. Doshi-Velez, Y. Ge, and I. Kohane, “Comorbidity clusters in autism spectrum disorders: an electronic health record time-series analysis”, *Pediatrics*, vol. 133, no. 1, e54–e63, 2014.

[189] B. E. Duffy and N. K. Chan, “‘‘you never really know who’s looking”: Imagined surveillance across social media platforms”, *New Media & Society*, vol. 21, no. 1, pp. 119–138, 2019.

[190] J. C. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters”, 1973.

[191] O. J. Dunn and V. Clark, “Comparison of tests of the equality of dependent correlation coefficients”, *Journal of the American Statistical Association*, vol. 66, no. 336, pp. 904–908, 1971.

[192] C. S. Dweck, “Mindsets: Developing talent through a growth mindset”, *Olympic Coach*, vol. 21, no. 1, pp. 4–7, 2009.

[193] K. Ehrlich and N. Shami, “Microblogging inside and outside the workplace”, in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 4, 2010.

[194] J. C. Eichstaedt and A. C. Weidman, “Tracking fluctuations in psychological states using social media language: A case study of weekly emotion”, *European Journal of Personality*, vol. 34, no. 5, pp. 845–858, 2020.

[195] D. Eisenberg, M. F. Downs, E. Golberstein, and K. Zivin, “Stigma and help seeking for mental health among college students”, *Medical Care Research and Review*, vol. 66, no. 5, pp. 522–541, 2009.

[196] D. Eisenberg, E. Golberstein, and S. E. Gollust, “Help-seeking and access to mental health care in a university student population”, *Medical care*, vol. 45, no. 7, pp. 594–601, 2007.

[197] D. Eisenberg, J. Hunt, and N. Speer, “Help seeking for mental health on college campuses: Review of evidence and next steps for research and practice”, *Harvard review of psychiatry*, 2012.

[198] R. Eisenberger, P. Fasolo, and V. Davis-LaMastro, “Perceived organizational support and employee diligence, commitment, and innovation.”, *Journal of applied psychology*, vol. 75, no. 1, p. 51, 1990.

[199] J. Elith, J. R. Leathwick, and T. Hastie, “A working guide to boosted regression trees”, *Journal of Animal Ecology*, 2008.

[200] N. B. Ellison, C. Steinfield, and C. Lampe, “The benefits of Facebook “friends”: Social capital and college studentsúse of online social network sites”, *Journal of Computer-Mediated Communication*, vol. 12, no. 4, pp. 1143–1168, 2007.

[201] C. Elsden, A. C. Durrant, and D. S. Kirk, “It’s just my history isn’t it? understanding smart journaling practices”, in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 2819–2831.

[202] A. Erez, V. F. Misangyi, D. E. Johnson, M. A. LePine, and K. C. Halverson, “Stirring the hearts of followers: charismatic leadership as the transferal of affect.”, *Journal of applied psychology*, vol. 93, no. 3, p. 602, 2008.

[203] S. K. Ernala, M. L. Birnbaum, K. A. Candan, A. F. Rizvi, W. A. Sterling, J. M. Kane, and M. De Choudhury, “Methodological gaps in predicting mental health states from social media: triangulating diagnostic signals”, in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–16.

[204] S. K. Ernala, K. H. Kashiparekh, A. Bolous, A. Asra, J. M. Kane, M. L. Birnbaum, and M. De Choudhury, “A social media study on mental health status transitions surrounding psychiatric hospitalizations”, *Proceedings of the ACM on Human-Computer Interaction*, no. CSCW, 2021.

[205] S. K. Ernala, T. Labetoulle, F. Bane, M. L. Birnbaum, A. F. Rizvi, J. M. Kane, and M. De Choudhury, “Characterizing audience engagement and assessing its impact on social media disclosures of mental illnesses”, in *ICWSM*, 2018.

[206] S. K. Ernala, A. F. Rizvi, M. L. Birnbaum, J. M. Kane, and M. De Choudhury, “Linguistic markers indicating therapeutic outcomes of social media disclosures of schizophrenia”, *Proceedings of the ACM on Human-Computer Interaction*, vol. 1, no. CSCW, pp. 1–27, 2017.

[207] D. Estrin, “Small data, where n= me”, *Communications of the ACM*, vol. 57, no. 4, pp. 32–34, 2014.

[208] B. Everitt and A. Skrondal, *The Cambridge dictionary of statistics*. Cambridge University Press Cambridge, 2002, vol. 106.

[209] J. Fan *et al.*, “Relationships between Five-Factor Personality Model and Anxiety: The Effect of Conscientiousness on Anxiety”, *Open Journal of Social Sciences*, vol. 8, no. 08, p. 462, 2020.

[210] R. Farzan, J. M. DiMicco, D. R. Millen, C. Dugan, W. Geyer, and E. A. Brownholtz, “Results from deploying a participation incentive mechanism within the enterprise”, in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 2008, pp. 563–572.

[211] E. Fast, B. Chen, and M. S. Bernstein, “Empath: Understanding topic signals in large-scale text”, in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, 2016, pp. 4647–4657.

[212] D. Ferreira, V. Kostakos, and A. K. Dey, “AWARE: mobile context instrumentation framework”, *Frontiers in ICT*, vol. 2, p. 6, 2015.

[213] C. Fiesler, M. Dye, J. L. Feuston, C. Hiruncharoenvate, C. J. Hutto, S. Morrison, P. Khanipour Roshan, U. Pavalanathan, A. S. Bruckman, M. De Choudhury, *et al.*, “What (or who) is public? privacy settings and social media content sharing”, in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 2017, pp. 567–580.

[214] C. Fiesler and N. Proferes, “‘‘participant” perceptions of twitter research ethics”, *Social Media+ Society*, vol. 4, no. 1, p. 2 056 305 118 763 366, 2018.

[215] F. Fiordelisi and O. Ricci, “Corporate culture and CEO turnover”, *Journal of Corporate Finance*, vol. 28, pp. 66–82, 2014.

[216] M. Fishbein, “The role of theory in HIV prevention”, *AIDS care*, vol. 12, no. 3, pp. 273–278, 2000.

[217] M. Fishbein and I. Ajzen, “Understanding attitudes and predicting social behavior”, 1980.

[218] M. Fishbein and J. N. Cappella, “The role of theory in developing effective health communications”, *Journal of communication*, vol. 56, S1–S17, 2006.

[219] W. A. Fisher, J. D. Fisher, and J. Harman, “The information-motivation-behavioral skills model: A general social psychological approach to understanding and promot-

ing health behavior", *Social psychological foundations of health and illness*, vol. 22, pp. 82–106, 2003.

[220] S. M. Fitzhugh, C. B. Gibson, E. S. Spiro, and C. T. Butts, "Spatio-temporal filtering techniques for the detection of disaster-related communication", *Social Science Research*, vol. 59, pp. 137–154, 2016.

[221] E. B. Foa, G. Steketee, and B. O. Rothbaum, "Behavioral/cognitive conceptualizations of post-traumatic stress disorder", *Behavior therapy*, vol. 20, no. 2, pp. 155–176, 1989.

[222] Fortune, *Fortune 500*, fortune.com/fortune500/list, Accessed: 2019-03-27, 2019.

[223] ——, *fortune.com/best-companies/list/*, Accessed: 2019-03-27, 2019.

[224] J. H. Fowler and N. A. Christakis, "Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the framingham heart study", *Bmj*, vol. 337, 2008.

[225] N. S. Fox, J. S. Brennan, and S. T. Chasen, "Clinical estimation of fetal weight and the Hawthorne effect", *European Journal of Obstetrics & Gynecology and Reproductive Biology*, vol. 141, no. 2, pp. 111–114, 2008.

[226] S. Fox, P. E. Spector, A. Goh, K. Bruursema, and S. R. Kessler, "The deviant citizen: Measuring potential positive relations between counterproductive work behaviour and organizational citizenship behaviour", *Journal of Occupational and Organizational Psychology*, vol. 85, no. 1, pp. 199–220, 2012.

[227] S. Freud, *Inhibitions, symptoms and anxiety*. WW Norton & Company, 1977.

[228] ——, *Beyond the pleasure principle*. Penguin UK, 2003.

[229] B. S. Frey and M. Osterloh, *Successful management by motivation: Balancing intrinsic and extrinsic incentives*. Springer Science & Business Media, 2001.

[230] Y. Fried, H. A. Ben-David, R. B. Tiegs, N. Avital, and U. Yeverechyahu, "The interactive effect of role conflict and role ambiguity on job performance", *Journal of occupational and organizational psychology*, 1998.

[231] A. Furnham, A. Eracleous, and T. Chamorro-Premuzic, "Personality, motivation and job satisfaction: Hertzberg meets the Big Five", *Journal of managerial psychology*, vol. 24, no. 8, pp. 765–779, 2009.

[232] R. P. Gallagher, "National survey of college counseling centers 2014", 2015.

[233] G. Galster, “A stock/flow model of defining racially integrated neighborhoods”, *Journal of Urban Affairs*, vol. 20, no. 1, pp. 43–51, 1998.

[234] Garmin, *Garmin health api*, http://developer.garmin.com/health-api/overview/, Accessed: 2018-11-01, 2018.

[235] W. Geyer, C. Dugan, J. DiMicco, D. R. Millen, B. Brownholtz, and M. Muller, “Use and reuse of shared lists as a social content type”, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2008, pp. 1545–1554.

[236] S. Ghoshray, “Employer surveillance versus employee privacy: The new reality of social media and workplace privacy”, *N. Ky. L. Rev.*, vol. 40, p. 593, 2013.

[237] E. Gilbert, “Phrases that signal workplace hierarchy”, in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ACM, 2012, pp. 1037–1046.

[238] Gimbal, *Manager rest api*, https://docs.gimbal.com/rest.html, Accessed: 2018-11-01, 2018.

[239] M. Gjoreski, H. Gjoreski, M. Lutrek, and M. Gams, “Automatic detection of perceived stress in campus students using smartphones”, in *Intelligent Environments (IE), 2015 International Conference on*, IEEE, 2015, pp. 132–135.

[240] S. R. Glaser, S. Zamanou, and K. Hacker, “Measuring and interpreting organizational culture”, *Management communication quarterly*, vol. 1, no. 2, pp. 173–198, 1987.

[241] K. Glasgow, C. Fink, and J. Boyd-Graber, “‘‘our grief is unspeakable”: Automatically measuring the community impact of a tragedy”, in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, 2014.

[242] Glassdoor, *Glassdoor give to get policy*, help.glassdoor.com/article/Give-to-get-policy/en_US, Accessed: 2019-04-04, 2018.

[243] ——, *Glassdoor statistics*, expandedramblings.com/index.php/numbers-15-interesting-glassdoor-statistics, Accessed: 2019-03-22, 2018.

[244] ——, *Glassdoor community guidelines*, https://https://help.glassdoor.com/article/Community-Guidelines/en_US, Accessed: 2019-03-21, 2019.

[245] ——, *Glassdoor fraudulent reviews*, help.glassdoor.com/article/Fraudulent-reviews/en_US/, Accessed: 2019-04-03, 2019.

[246] A. M. Glenberg, M. M. Bradley, T. A. Kraus, and G. J. Renzaglia, “Studies of the long-term recency effect: Support for a contextually guided retrieval hypothesis.”,

Journal of Experimental Psychology: Learning, Memory, and Cognition, vol. 9, no. 2, p. 231, 1983.

[247] I. Gochman, “Arousal, attribution, and environmental stress”, *Stress and anxiety*, vol. 6, pp. 67–92, 1979.

[248] K. R. Gøeg, R. Cornet, and S. K. Andersen, “Clustering clinical models from local electronic health records based on semantic similarity”, *Journal of biomedical informatics*, vol. 54, pp. 294–304, 2015.

[249] E. Goffman, “The presentation of self in everyday life”, 1959.

[250] ——, *Forms of talk*. University of Pennsylvania Press, 1981.

[251] E. L. Goldberg and G. W. Comstock, “Epidemiology of life events: Frequency in general populations”, *American Journal of epidemiology*, 1980.

[252] S. A. Golder and M. W. Macy, “Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures”, *Science*, vol. 333, no. 6051, pp. 1878–1881, 2011.

[253] F. Gomez, *A Guide to the Depression, Anxiety and Stress Scale (DASS 21)*, 2016.

[254] J. L. Goodie, K. T. Larkin, and S. Schauss, “Validation of Polar heart rate monitor for assessing heart rate during physical and mental stress.”, *Journal of Psychophysiology*, vol. 14, no. 3, p. 159, 2000.

[255] Google, *Activity recognition api*, https://developers.google.com/location-context/activity-recognition/, Accessed: 2018-11-01, 2018.

[256] *Google bigquery*, bigquery.cloud.google.com, Accessed: 2017-10-27, 2017.

[257] M. Goos, A. Manning, and A. Salomons, “Explaining job polarization: the roles of technology, offshoring and institutions”, *Offshoring and Institutions (December 1, 2011)*, 2011.

[258] A. M. Grant, J. M. Berg, and D. M. Cable, “Job titles as identity badges: How self-reflective titles can reduce emotional exhaustion”, *Academy of Management journal*, vol. 57, no. 4, pp. 1201–1225, 2014.

[259] A. E. Green, Y. Li, D. Owen, and M. De Hoyos, “Inequalities in use of the Internet for job search: similarities and contrasts by economic status in Great Britain”, *Environment and Planning A*, vol. 44, no. 10, pp. 2344–2358, 2012.

[260] S. Greengard, “Digitally possessed”, *Communications of the ACM*, 2012.

[261] J. H. Greenhaus and S. Parasuraman, “Job performance attributions and career advancement prospects: An examination of gender and race effects”, *Organizational Behavior and Human Decision Processes*, vol. 55, no. 2, pp. 273–297, 1993.

[262] S. Greenwood, A. Perrin, and M. Duggan, *Demographics of Social Media Users in 2016*, pewinternet.org/2016/11/11/social-media-update-2016/, Accessed: 2017-02-12, 2016.

[263] ——, “Social media update 2016”, *Pew Research Center*, vol. 11, no. 2, pp. 1–18, 2016.

[264] M. A. Griffin, A. Neal, and S. K. Parker, “A new model of work role performance: Positive behavior in uncertain and interdependent contexts”, *Academy of management journal*, vol. 50, no. 2, pp. 327–347, 2007.

[265] P. Grimm, “Social desirability bias”, *Wiley international encyclopedia of marketing*, 2010.

[266] U. Grömping, “Variable importance assessment in regression: linear regression versus random forest”, *The American Statistician*, vol. 63, no. 4, pp. 308–319, 2009.

[267] J. J. Gross and R. A. Thompson, “Emotion regulation: Conceptual foundations.”, 2007.

[268] T. Grover and G. Mark, “Digital footprints: Predicting personality from temporal patterns of technology use”, in *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 2017, pp. 41–44.

[269] A. Grünerbl, A. Muaremi, V. Osmani, G. Bahle, S. Oehler, G. Tröster, O. Mayora, C. Haring, and P. Lukowicz, “Smartphone-based recognition of states and state changes in bipolar disorder patients”, *IEEE JBHI*, 2015.

[270] X. S. Gu and P. R. Rosenbaum, “Comparison of multivariate matching methods: Structures, distances, and algorithms”, *Journal of Computational and Graphical Statistics*, vol. 2, no. 4, pp. 405–420, 1993.

[271] T. Guberek, A. McDonald, S. Simioni, A. H. Mhaidli, K. Toyama, and F. Schaub, “Keeping a low profile? technology, risk and privacy among undocumented immigrants”, in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018.

[272] B. Guerin, “Social facilitation”, *The Corsini encyclopedia of psychology*, pp. 1–2, 2010.

[273] M.-D. Guillamon, A.-M. Rios, B. Gesuele, and C. Metallo, "Factors influencing social media use in local governments: The case of italy and spain", *Government Information Quarterly*, vol. 33, no. 3, pp. 460–471, 2016.

[274] J. Guillory and J. T. Hancock, "The effect of Linkedin on deception in resumes", *Cyberpsychology, Behavior, and Social Networking*, 2012.

[275] J. Guillory, J. Spiegel, M. Drislane, B. Weiss, W. Donner, and J. Hancock, "Upset now? emotion contagion in distributed groups", in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2011, pp. 745–748.

[276] J. J. Gumperz and D. H. Hymes, *Directions in sociolinguistics: The ethnography of communication*. B. Blackwell, 1986.

[277] S. C. Guntuku, A. Buffone, K. Jaidka, J. C. Eichstaedt, and L. H. Ungar, "Understanding and measuring psychological stress using social media", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 13, 2019, pp. 214–225.

[278] S. C. Guntuku, J. R. Ramsay, R. M. Merchant, and L. H. Ungar, "Language of ADHD in adults on social media", *Journal of attention disorders*, 2019.

[279] S. C. Guntuku, R. Schneider, A. Pelullo, J. Young, V. Wong, L. Ungar, D. Polsky, K. G. Volpp, and R. Merchant, "Studying expressions of loneliness in individuals using twitter: An observational study", *BMJ open*, vol. 9, no. 11, 2019.

[280] S. C. Guntuku, H. A. Schwartz, A. Kashyap, J. S. Gaulton, D. C. Stokes, D. A. Asch, L. H. Ungar, and R. M. Merchant, "Variability in language used on social media prior to hospital visits", *Scientific reports*, vol. 10, no. 1, pp. 1–9, 2020.

[281] A. K. Gupta and V. Govindarajan, "Cultivating a global mindset", *Academy of Management Perspectives*, vol. 16, no. 1, pp. 116–126, 2002.

[282] N. Haferkamp, S. C. Eimler, A.-M. Papadakis, and J. V. Kruck, "Men are from Mars, women are from Venus? Examining gender differences in self-presentation on social networking sites", *Cyberpsychology, Behavior, and Social Networking*, vol. 15, no. 2, pp. 91–98, 2012.

[283] O. L. Haimson, A. E. Bowser, E. F. Melcer, and E. F. Churchill, "Online inspiration and exploration for identity reinvention", in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 3809–3818.

[284] O. L. Haimson, J. R. Brubaker, L. Dombrowski, and G. R. Hayes, "Disclosure, stress, and support during gender transition on facebook", in *Proceedings of the*

18th ACM Conference on Computer Supported Cooperative Work & Social Computing, 2015, pp. 1176–1190.

[285] O. L. Haimson, A. J. Carter, S. Corvite, B. Wheeler, L. Wang, T. Liu, and A. Lige, "The major life events taxonomy: Social readjustment, social media information sharing, and online network separation during times of life transition", *Journal of the Association for Information Science and Technology*, 2021.

[286] O. L. Haimson and T. C. Veinot, "Coming out to doctors, coming out to "everyone": Understanding the average sequence of transgender identity disclosures using social media data", *Transgender Health*, 2020.

[287] J. A. Hall and N. Pennington, "Self-monitoring, honesty, and cue use on facebook: The relationship with user extraversion and conscientiousness", *Computers in Human Behavior*, vol. 29, no. 4, pp. 1556–1564, 2013.

[288] J. Handl, J. Knowles, and D. B. Kell, "Computational cluster validation in post-genomic data analysis", *Bioinformatics*, vol. 21, no. 15, pp. 3201–3212, 2005.

[289] P. D. Harvey, B. R. Greenberg, and M. R. Serper, "The affective lability scales: development, reliability, and validity", *Journal of clinical psychology*, vol. 45, no. 5, pp. 786–793, 1989.

[290] Q. He, C. A. Glas, M. Kosinski, D. J. Stillwell, and B. P. Veldkamp, "Predicting self-monitoring skills using textual posts on facebook", *Computers in Human Behavior*, vol. 33, pp. 69–78, 2014.

[291] J. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors", *Intelligent Transportation Systems, IEEE Transactions on*, vol. 6, no. 2, pp. 156–166, 2005.

[292] H. Heidari, V. Nanda, and K. Gummadi, "On the Long-term Impact of Algorithmic Decision Policies: Effort Unfairness and Feature Segregation through Social Learning", in *International Conference on Machine Learning*, 2019, pp. 2692–2701.

[293] H. Heidemeier and K. Moser, "Self–other agreement in job performance ratings: A meta-analytic test of a process model.", *Journal of Applied Psychology*, vol. 94, no. 2, p. 353, 2009.

[294] E. Hekler, J. A. Tiro, C. M. Hunter, and C. Nebeker, "Precision health: The role of the social and behavioral sciences in advancing the vision", *Annals of Behavioral Medicine*, 2020.

[295] J. M. Hektner, J. A. Schmidt, and M. Csikszentmihalyi, *Experience sampling method: Measuring the quality of everyday life*. Sage, 2007.

[296] J. Hellhammer and M. Schubert, “The physiological response to Trier Social Stress Test relates to subjective measures of stress during but not before or after the test”, *Psychoneuroendocrinology*, vol. 37, no. 1, pp. 119–124, 2012.

[297] R. K. Henson and J. K. Roberts, “Use of exploratory factor analysis in published research: Common errors and some comment on improved practice”, *Educ. Psychol. Meas.*, 2006.

[298] P. L. Hewitt, G. L. Flett, and S. W. Mosher, “The Perceived Stress Scale: Factor structure and relation to depression symptoms in a psychiatric sample”, *Journal of Psychopathology and Behavioral Assessment*, vol. 14, no. 3, pp. 247–257, 1992.

[299] L. Hickman, K. Saha, M. De Choudhury, and L. Tay, “Automated tracking of components of job satisfaction via text mining of twitter data”, in *ML Symposium, SIOP*, 2019.

[300] J. M. Hilbe, *Negative binomial regression*. Cambridge University Press, 2011.

[301] K. Hill, “How target figured out a teen girl was pregnant before her father did”, *Forbes, Inc*, 2012.

[302] J. K. Hirsch and J. B. Ellis, “Differences in life stress and reasons for living among college suicide ideators and non-ideators.”, *College student journal*, 1996.

[303] G. Hofstede, *Culture’s consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage publications, 2001.

[304] G. Hofstede, B. Neuijen, D. D. Ohayv, and G. Sanders, “Measuring organizational cultures: A qualitative and quantitative study across twenty cases”, *Administrative science quarterly*, pp. 286–316, 1990.

[305] B. Hogan, “The presentation of self in the age of social media: Distinguishing performances and exhibitions online”, *Bulletin of Science, Technology & Society*, vol. 30, no. 6, pp. 377–386, 2010.

[306] T. C. Holdeman, “Invisible wounds of war: Psychological and cognitive injuries, their consequences, and services to assist recovery”, *Psychiatric Services*, vol. 60, no. 2, pp. 273–273, 2009.

[307] P. W. Holland, “Statistics and causal inference”, *Journal of the American statistical Association*, vol. 81, no. 396, pp. 945–960, 1986.

[308] E. E. Hollenbaugh and A. L. Ferris, “Facebook self-disclosure: Examining the role of traits, social cohesion, and motives”, *Computers in Human Behavior*, vol. 30, pp. 50–58, 2014.

[309] M. C. Howard and M. E. Hoffman, "Variable-centered, person-centered, and person-specific approaches: where theory meets the method", *Organizational Research Methods*, vol. 21, no. 4, pp. 846–876, 2018.

[310] H.-F. Hsieh and S. E. Shannon, "Three approaches to qualitative content analysis", *Qualitative health research*, vol. 15, no. 9, pp. 1277–1288, 2005.

[311] Y. Huang, Y. Tang, and Y. Wang, "Emotion map: A location-based mobile social system for improving emotion awareness and regulation", in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, ACM, 2015, pp. 130–142.

[312] R. Huff-Eibl, J. F. Voyles, and M. M. Brewer, "Competency-based hiring, job description, and performance goals: The value of an integrated system", *Journal of Library Administration*, vol. 51, no. 7-8, pp. 673–691, 2011.

[313] W. N. I. W. Husina, A. Santosa, H. M. Ramosa, and M. S. Nordinb, "Crystallized intelligence or fluid intelligence factor?", *Procedia-Social and Behavioral Sciences*, vol. 97, pp. 214–223, 2013.

[314] B. Hutchinson and M. Mitchell, "50 years of test (un) fairness: Lessons for machine learning", in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2019, pp. 49–58.

[315] C. Hutto and E. Gilbert, *Vader: A parsimonious rule-based model for sentiment analysis of social media text. icwsm proceedings*, 2014.

[316] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy", *International journal of forecasting*, 2006.

[317] G. W. Imbens and D. B. Rubin, *Causal inference in statistics, social, and biomedical sciences*. Cambridge, 2015.

[318] R. Ingrassia, "Appearance or reality? Monitoring of employer branding in public network space: the Glassdoor case", 2017.

[319] M. Irvine, *Google releases life events targeting to everyone!*, https://www.wordstream.com/blog/ws/2017/11/29/adwords-life-events-targeting, Accessed: 2020-09-08, 2018.

[320] S. E. Jackson and R. S. Schuler, "A meta-analysis and conceptual critique of research on role ambiguity and role conflict in work settings", *Organizational behavior and human decision processes*, vol. 36, no. 1, pp. 16–78, 1985.

[321] W. S. Jacobson and S. H. Tufts, “To post or not to post: Employee rights and social media”, *Review of public personnel administration*, vol. 33, no. 1, pp. 84–107, 2013.

[322] M. Jacovi, I. Guy, S. Kremer-Davidson, S. Porat, and N. Aizenbud-Reshef, “The perception of others: inferring reputation from social media in the enterprise”, in *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, ACM, 2014, pp. 756–766.

[323] S. Jahng, P. K. Wood, and T. J. Trull, “Analysis of affective instability in ecological momentary assessment: Indices using successive difference and group comparison via multilevel modeling.”, *Psychological methods*, vol. 13, no. 4, p. 354, 2008.

[324] K. Jaidka, S. Giorgi, H. A. Schwartz, M. L. Kern, L. H. Ungar, and J. C. Eichstaedt, “Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods”, *Proceedings of the National Academy of Sciences*, vol. 117, no. 19, pp. 10 165–10 171, 2020.

[325] J. W. James and R. Friedman, *The Grief Recovery Handbook, 20th Anniversary Expanded Edition: The Action Program for Moving Beyond Death, Divorce, and Other Losses including Health, Career, and Faith*. Harper Collins, 2009.

[326] K. A. Jehn, “A qualitative analysis of conflict types and dimensions in organizational groups”, *Administrative science quarterly*, pp. 530–557, 1997.

[327] S. Jhaver, J. Cranshaw, and S. Counts, “Measuring professional skill development in us cities using internet search queries”, in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 13, 2019, pp. 267–277.

[328] S. Jhaver, S. Ghoshal, A. Bruckman, and E. Gilbert, “Online harassment and content moderation: The case of blocklists”, *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 25, no. 2, p. 12, 2018.

[329] J. V. Johnson and J. Lipscomb, “Long working hours, occupational health and the changing nature of work organization”, *American journal of industrial medicine*, vol. 49, no. 11, pp. 921–929, 2006.

[330] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.

[331] J. Jouhki, E. Lauk, M. Penttinen, N. Sormanen, and T. Uskali, “Facebook’s emotional contagion experiment as a challenge to research ethics”, *Media and Communication*, vol. 4, 2016.

[332] D. Jurafsky and H. James, “Speech and language processing an introduction to natural language processing, computational linguistics, and speech”, 2000.

[333] R. Jurdak, K. Zhao, J. Liu, M. AbouJaoude, M. Cameron, and D. Newth, "Understanding human mobility from Twitter", *PloS one*, vol. 10, no. 7, e0131469, 2015.

[334] P. Juszczak, D. Tax, and R. P. Duin, "Feature scaling in support vector data description", in *Proc. ASCI*, Citeseer, 2002, pp. 95–102.

[335] R. L. Kahn and P. Byosiere, "Stress in organizations.", 1992.

[336] R. L. Kahn, D. M. Wolfe, R. P. Quinn, J. D. Snoek, and R. A. Rosenthal, "Organizational stress: Studies in role conflict and ambiguity.", 1964.

[337] S. Kairam, M. Brzozowski, D. Huffaker, and E. Chi, "Talking in circles: selective sharing in google+", in *Proceedings of the SIGCHI conference on human factors in computing systems*, ACM, 2012, pp. 1065–1074.

[338] S. R. Kairam, D. J. Wang, and J. Leskovec, "The life and death of online groups: Predicting group growth and longevity", in *Proceedings of the fifth ACM international conference on Web search and data mining*, 2012, pp. 673–682.

[339] C. C. S. Kam and J. P. Meyer, "How careless responding and acquiescence response bias can influence construct dimensionality: The case of job satisfaction", *Organizational Research Methods*, vol. 18, no. 3, pp. 512–541, 2015.

[340] R. Kang, S. Brown, and S. Kiesler, "Why do people seek anonymity on the internet? informing policy and design", in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 2657–2666.

[341] S.-G. Kang, J. M. Kang, K.-P. Ko, S.-C. Park, S. Mariani, and J. Weng, "Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers", *Journal of psychosomatic research*, vol. 97, pp. 38–44, 2017.

[342] A. E. Kazdin, "Observer effects: Reactivity of direct observation.", *New Directions for Methodology of Social & Behavioral Science*, 1982.

[343] ——, "Evidence-based treatment research: Advances, limitations, and next steps.", *American Psychologist*, vol. 66, no. 8, p. 685, 2011.

[344] E. A. Kensinger and S. Corkin, "Two routes to emotional memory: Distinct neural processes for valence and arousal", *Proceedings of the National Academy of Sciences*, vol. 101, no. 9, pp. 3310–3315, 2004.

[345] L. A. King and D. W. King, "Role conflict and role ambiguity: A critical assessment of construct validity.", *Psychological Bulletin*, vol. 107, no. 1, p. 48, 1990.

[346] E. Kıcıman, S. Counts, and M. Gasser, “Using Longitudinal Social Media Analysis to Understand the Effects of Early College Alcohol Use.”, in *ICWSM*, 2018, pp. 171–180.

[347] K. Klein and A. Boals, “Expressive writing can increase working memory capacity.”, *Journal of experimental psychology: General*, vol. 130, no. 3, p. 520, 2001.

[348] R. Kline, “Resisting consumer technology in rural america: The telephone and electrification”, *How users matter: The co-construction of users and technology*, pp. 51–66, 2003.

[349] D. E. Klingner, “When the traditional job description is not enough”, *Personnel journal*, 1979.

[350] D. K. Knight, C. Crutsinger, and H. Kim, “The impact of retail work experience, career expectation, and job satisfaction on retail career intention”, *Clothing and Textiles Research Journal*, vol. 24, no. 1, pp. 1–14, 2006.

[351] T. Koch, C. Gerber, and J. J. de Klerk, “The impact of social media on recruitment: Are you LinkedIn?”, *SA Journal of Human Resource Management*, vol. 16, no. 1, pp. 1–14, 2018.

[352] R. Kohavi *et al.*, “A study of cross-validation and bootstrap for accuracy estimation and model selection”, in *Ijcai*, Montreal, Canada, vol. 14, 1995, pp. 1137–1145.

[353] P. Kollock, “The economies ol online cooperation”, *Communities in cyberspace*, vol. 220, 1999.

[354] K. Kolmes, “Social media in the future of professional psychology.”, *Professional Psychology: Research and Practice*, vol. 43, no. 6, p. 606, 2012.

[355] M. Komarraju, S. J. Karau, R. R. Schmeck, and A. Avdic, “The Big Five personality traits, learning styles, and academic achievement”, *Personality and individual differences*, vol. 51, no. 4, pp. 472–477, 2011.

[356] M. Kosinski, D. Stillwell, and T. Graepel, “Private traits and attributes are predictable from digital records of human behavior”, *Proceedings of the national academy of sciences*, vol. 110, no. 15, pp. 5802–5805, 2013.

[357] A. D. Kramer, J. E. Guillory, and J. T. Hancock, “Experimental evidence of massive-scale emotional contagion through social networks”, *Proceedings of the National Academy of Sciences*, vol. 111, no. 24, pp. 8788–8790, 2014.

[358] M. Kumar, M. Dredze, G. Coppersmith, and M. De Choudhury, “Detecting changes in suicide content manifested in social media following celebrity suicides”, in *ACM HT*, 2015.

[359] J. Kuoppala, A. Lamminpää, J. Liira, and H. Vainio, “Leadership, job well-being, and health effects—a systematic review and a meta-analysis”, *Journal of occupational and environmental medicine*, vol. 50, no. 8, 2008.

[360] M. Kusserow, O. Amft, and G. Troster, “Modeling arousal phases in daily living using wearable sensors”, *Affective Computing, IEEE Transactions on*, vol. 4, no. 1, pp. 93–105, 2013.

[361] N. Ladany and M. L. Friedlander, “The relationship between the supervisory working alliance and trainees’ experience of role conflict and role ambiguity”, *Counselor Education and supervision*, vol. 34, no. 3, pp. 220–231, 1995.

[362] H.-M. Lai and T. T. Chen, “Knowledge sharing in interest online communities: A comparison of posters and lurkers”, *Computers in Human Behavior*, vol. 35, pp. 295–306, 2014.

[363] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, “A survey of mobile phone sensing”, *IEEE Communications magazine*, vol. 48, no. 9, 2010.

[364] B. Laursen, W. Furman, and K. S. Mooney, “Predicting interpersonal competence and self-worth from adolescent relationships and relationship networks: Variable-centered and person-centered perspectives”, *Merrill-Palmer Quarterly (1982-)*, 2006.

[365] R. S. Lazarus, “Psychological stress in the workplace”, *Occupational stress: A handbook*, vol. 1, pp. 3–14, 1995.

[366] D. Lazer, R. Kennedy, G. King, and A. Vespignani, “The parable of Google Flu: traps in big data analysis”, *Science*, vol. 343, no. 6176, pp. 1203–1205, 2014.

[367] M. D. LeCompte and J. P. Goetz, “Problems of reliability and validity in ethnographic research”, *Review of educational research*, vol. 52, no. 1, pp. 31–60, 1982.

[368] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning”, *nature*, 2015.

[369] C. S. Lee and L. Ma, “News sharing in social media: The effect of gratifications and prior experience”, *Computers in human behavior*, vol. 28, no. 2, pp. 331–339, 2012.

[370] J. A. Lee, C. Efstratiou, and L. Bai, "OSN mood tracking: exploring the use of online social network activity as an indicator of mood changes", in *Ubicomp: Adjunct*, 2016.

[371] J. Lee and J. Kang, "A study on job satisfaction factors in retention and turnover groups using dominance analysis and lda topic modeling with employee reviews on glassdoor.com", 2017.

[372] J. A. LePine, J. A. Colquitt, and A. Erez, "Adaptability to changing task contexts: Effects of general cognitive ability, conscientiousness, and openness to experience", *Personnel psychology*, vol. 53, no. 3, pp. 563–593, 2000.

[373] P. LeViness, C. Bershad, and K. Gorman, "The association for university and college counseling center directors annual survey", *Retrieved fro m https://www. aucccd. org/assets/documents/Governance/2017% 20aucccd% 20surveypublicapr26. pdf*, 2017.

[374] B. Li and A. Sano, "Extraction and Interpretation of Deep Autoencoder-based Temporal Features from Wearables for Forecasting Personalized Mood, Health, and Stress", *PACM IMWUT*, 2020.

[375] Z. Li, K. Zhang, Y. Xie, F. Yu, and X. Wang, "Knowing your enemy: Understanding and detecting malicious web advertising", in *Proceedings of the 2012 ACM conference on Computer and communications security*, 2012, pp. 674–686.

[376] P. Liao, K. Greenewald, P. Klasnja, and S. Murphy, "Personalized HeartSteps: A Reinforcement Learning Algorithm for Optimizing Physical Activity", *PACM IMWUT*, 2020.

[377] Y.-R. Lin and D. Margolin, "The ripple of fear, sympathy and solidarity during the Boston bombings", *EPJ Data Science*, vol. 3, no. 1, p. 31, 2014.

[378] W. Little, R. McGivern, and N. Kerins, *Introduction to Sociology-2nd Canadian Edition*. BC Campus, 2016.

[379] C. H. Liu, C. Stevens, S. H. Wong, M. Yasui, and J. A. Chen, "The prevalence and predictors of mental health diagnoses and suicide among US college students: Implications for addressing disparities in service use", *Depression and anxiety*, vol. 36, no. 1, pp. 8–17, 2019.

[380] J. Liu, E. R. Weitzman, and R. Chunara, "Assessing behavior stage progression from social media data", in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 2017, pp. 1320–1333.

[381] S. Liu, M. Zhu, and S. D. Young, “Monitoring freshman college experience through content analysis of tweets: observational study”, *JMIR public health and surveillance*, vol. 4, no. 1, e5, 2018.

[382] H. Lu, D. Frauendorfer, M. Rabbi, M. S. Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury, “Stresssense: Detecting stress in unconstrained acoustic environments using smartphones”, in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ACM, 2012, pp. 351–360.

[383] B. Lubin, S. H. Gardener, and A. Roth, “Mood and somatic symptoms during pregnancy.”, *Psychosomatic Medicine*, vol. 37, no. 2, pp. 136–146, 1975.

[384] M. Luca and G. Zervas, “Fake it till you make it: Reputation, competition, and Yelp review fraud”, *Management Science*, vol. 62, no. 12, pp. 3412–3427, 2016.

[385] R. E. Lucas, E. Diener, A. Grob, E. M. Suh, and L. Shao, “Cross-cultural evidence for the fundamental features of extraversion.”, *Journal of personality and social psychology*, vol. 79, no. 3, p. 452, 2000.

[386] R. E. Lucas and F. Fujita, “Factors influencing the relation between extraversion and pleasant affect.”, *Journal of personality and social psychology*, vol. 79, no. 6, p. 1039, 2000.

[387] L. Luo, “Work motivation, job stress and employees’ well-being”, *Journal of applied management studies*, vol. 8, 1999.

[388] S. Lyubomirsky, “Why are some people happier than others? The role of cognitive and motivational processes in well-being.”, *American psychologist*, vol. 56, no. 3, p. 239, 2001.

[389] X. Ma, J. Hancock, and M. Naaman, “Anonymity, intimacy and self-disclosure in social media”, in *Proceedings of the 2016 CHI conference on human factors in computing systems*, ACM, 2016, pp. 3857–3869.

[390] T. J. Madden, P. S. Ellen, and I. Ajzen, “A comparison of the theory of planned behavior and the theory of reasoned action”, *Personality and social psychology Bulletin*, vol. 18, no. 1, pp. 3–9, 1992.

[391] D. Magnusson, *The logic and implications of a person-oriented approach.* Sage Publications, Inc, 1998.

[392] A. M. Manago, T. Taylor, and P. M. Greenfield, “Me and my 400 friends: the anatomy of college students’ Facebook networks, their communication patterns, and well-being.”, *Developmental psychology*, 2012.

[393] C. Manning, M. Surdeanu, J. Bauer, J. Finkel, S. Bethard, and D. McClosky, "The Stanford CoreNLP natural language processing toolkit", in *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*, 2014, pp. 55–60.

[394] L. Manovich, "Trending: The promises and the challenges of big social data", *Debates in the digital humanities*, vol. 2, pp. 460–475, 2011.

[395] J. Manyika, S. Lund, M. Chui, J. Bughin, J. Woetzel, P. Batra, R. Ko, and S. Sanghvi, "Jobs lost, jobs gained: Workforce transitions in a time of automation", *McKinsey Global Institute*, 2017.

[396] P. Maquet, J.-M. Péters, J. Aerts, G. Delfiore, C. Degueldre, A. Luxen, and G. Franck, "Functional neuroanatomy of human rapid-eye-movement sleep and dreaming", *Nature*, vol. 383, no. 6596, pp. 163–166, 1996.

[397] G. Mark, M. Bagdouri, L. Palen, J. Martin, B. Al-Ani, and K. Anderson, "Blogs as a collective war diary", in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, 2012, pp. 37–46.

[398] G. Mark, S. T. Iqbal, M. Czerwinski, and P. Johns, "Bored mondays and focused afternoons: The rhythm of attention and online activity in the workplace", in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2014, pp. 3025–3034.

[399] G. Mark, Y. Wang, M. Niiya, and S. Reich, "Sleep Debt in Student Life: Online Attention Focus, Facebook, and Mood", in *Proc. 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 5517–5528.

[400] G. J. Martinez, S. M. Mattingly, J. Young, L. Faust, A. K. Dey, A. T. Campbell, M. De Choudhury, S. Mirjafari, S. Nepal, P. Robles-Granda, *et al.*, "Improved Sleep Detection Through The Fusion of Phone Agent and Wearable Data Streams", in *WristSense 2020*, 2020.

[401] S. Marwaha, M. R. Broome, P. E. Bebbington, E. Kuipers, and D. Freeman, "Mood instability and psychosis: analyses of British national survey data", *Schizophrenia bulletin*, vol. 40, no. 2, pp. 269–277, 2014.

[402] A. E. Marwick, "Entrepreneurial Subjects: Venturing from Alley to Valley.", *International Journal of Communication (19328036)*, vol. 11, 2017.

[403] A. E. Marwick and D. Boyd, "I tweet honestly, i tweet passionately: Twitter users, context collapse, and the imagined audience", *New media & society*, vol. 13, no. 1, pp. 114–133, 2011.

[404] M. Massimi and R. M. Baecker, “A death in the family: Opportunities for designing technologies for the bereaved”, in *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 2010, pp. 1821–1830.

[405] M. Masuda and T. H. Holmes, “Life events: Perceptions and frequencies”, *Psychosomatic medicine*, vol. 40, no. 3, pp. 236–261, 1978.

[406] S. M. Mattingly, J. M. Gregg, P. Audia, A. E. Bayraktaroglu, A. T. Campbell, N. V. Chawla, V. D. Swain, M. De Choudhury, S. K. D’Mello, A. K. Dey, *et al.*, “The Tesserae Project: Large-Scale, Longitudinal, In Situ, Multimodal Sensing of Information Workers”, 2019.

[407] I. L. McCann and L. A. Pearlman, *Psychological trauma and the adult survivor: Theory, therapy, and transformation*, 21. Psychology Press, 1990.

[408] D. McDowall, R. McCleary, and B. J. Bartos, *Interrupted time series analysis*. Oxford University Press, 2019.

[409] A. C. McFARLANE, “The long-term costs of traumatic stress: intertwined physical and psychological consequences”, *World Psychiatry*, vol. 9, no. 1, pp. 3–10, 2010.

[410] J. E. McGrath, “Stress and behavior in organization. Handbook of industrial and organizational psychology”, *Chicago: RandMcnall*, 1976.

[411] D. E. McNabb and F. T. Sepic, “Culture, climate, and total quality management: Measuring readiness for change”, *Public Productivity & Management Review*, pp. 369–385, 1995.

[412] A. Mehrotra, V. Pejovic, and M. Musolesi, “SenSocial: a middleware for integrating online social networks and mobile sensing data streams”, in *Proceedings of the 15th International Middleware Conference*, ACM, 2014, pp. 205–216.

[413] R. Mehrotra, S. Sanner, W. Buntine, and L. Xie, “Improving lda topic models for microblogs via tweet pooling and automatic labeling”, in *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, 2013, pp. 889–892.

[414] R. K. Merton, “The Matthew effect in science: The reward and communication systems of science are considered”, *Science*, vol. 159, no. 3810, pp. 56–63, 1968.

[415] E. Meyer, *Inadvertent Algorithmic Cruelty*, https://meyerweb.com/eric/thoughts/2014/12/24/inadvertent-algorithmic-cruelty/, Accessed: 2020-09-08, 2014.

[416] J. Middleton, W. Buboltz, and B. Sopon, “The relationship between psychological reactance and emotional intelligence”, *The Social Science Journal*, vol. 52, no. 4, pp. 542–549, 2015.

[417] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space”, *arXiv preprint arXiv:1301.3781*, 2013.

[418] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality”, in *Advances in neural information processing systems*, 2013, pp. 3111–3119.

[419] J. Miles, “Tolerance and variance inflation factor”, *Wiley StatsRef: Statistics Reference Online*, 2014.

[420] C. H. Miller, M. Burgoon, J. R. Grandpre, and E. M. Alvaro, “Identifying principal risk factors for the initiation of adolescent smoking behaviors: The significance of psychological reactance”, *Health communication*, vol. 19, no. 3, pp. 241–252, 2006.

[421] H. Miller, “The presentation of self in electronic life: Goffman on the Internet”, in *Embodied knowledge and virtual space conference*, vol. 9, 1995.

[422] S. Mirjafari, K. Masaba, T. Grover, W. Wang, P. Audia, A. T. Campbell, N. V. Chawla, M. De Choudhury, *et al.*, “Differentiating Higher and Lower Job Performers in the Workplace using Mobile Sensing”, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1–23, 2019.

[423] S. Mirjafari, K. Masaba, T. Grover, W. Wang, P. Audia, A. T. Campbell, N. V. Chawla, V. D. Swain, M. D. Choudhury, A. K. Dey, and et al., “Differentiating Higher and Lower Job Performers in the Workplace Using Mobile Sensing”, *Proc. ACM IMWUT*, 2019.

[424] T. Mitra and E. Gilbert, “Have you heard?: How gossip flows through workplace email”, in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 6, 2012.

[425] T. Mitra, M. Muller, N. S. Shami, A. Golestani, and M. Masli, “Spread of Employee Engagement in a Large Organizational Network: A Longitudinal Analysis”, *Proceedings of the ACM on Human-Computer Interaction*, vol. 1, no. CSCW, p. 81, 2017.

[426] A. Mohan, Z. Chen, and K. Weinberger, “Web-search ranking with initialized gradient boosted regression trees”, in *Proceedings of the learning to rank challenge*, 2011, pp. 77–89.

[427] T. Monahan and J. A. Fisher, "Benefits of 'observer effects': lessons from the field", *Qualitative research*, vol. 10, no. 3, pp. 357–376, 2010.

[428] M. A. Moreno, D. A. Christakis, K. G. Egan, L. N. Brockman, and T. Becker, "Associations between displayed alcohol references on facebook and problem drinking among college students", *Archives of Pediatrics & Adolescent Medicine*, vol. 166, no. 2, pp. 157–163, 2011.

[429] M. A. Moreno, L. A. Jelenchick, K. G. Egan, E. Cox, H. Young, K. E. Gannon, and T. Becker, "Feeling bad on Facebook: depression disclosures by college students on a social networking site", *Depress. Anxiety*, 2011.

[430] E. M. Morgan, C. Snelson, and P. Elison-Bowers, "Image and video disclosure of substance use on social media websites", *Computers in Human Behavior*, vol. 26, no. 6, pp. 1405–1411, 2010.

[431] E. W. Morrison and F. J. Milliken, "Organizational silence: A barrier to change and development in a pluralistic world", *Academy of Management review*, vol. 25, no. 4, pp. 706–725, 2000.

[432] J. J. Muehlenkamp, S. Marrone, J. S. Gray, and D. L. Brown, "A college suicide prevention model for american indian students.", *Professional Psychology: Research and Practice*, vol. 40, no. 2, p. 134, 2009.

[433] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, "What yelp fake review filter might be doing?", in *Seventh international AAAI conference on weblogs and social media*, 2013.

[434] D. A. Murphey, "Presenting community-level data in an "outcomes and indicators" framework: Lessons from vermont's experience", *Public Administration Review*, pp. 76–82, 1999.

[435] D. A. Nadler, G. D. Jenkins, C. Cammann, and E. E. Lawler, "The Michigan organizational assessment package: Progress report II", *Ann Arbor: Institute for Social Research, University of Michigan*, 1975.

[436] J. S. Nairne, S. R. Thompson, and J. N. Pandeirada, "Adaptive memory: Survival processing enhances retention.", *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 33, no. 2, p. 263, 2007.

[437] M. W. Newman, D. Lauterbach, S. A. Munson, P. Resnick, and M. E. Morris, "It's not that i don't have problems, i'm just not putting them on facebook: Challenges and opportunities in using online social networks for health", in *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, 2011, pp. 341–350.

[438] U. News, *News*, usnews.com, Accessed: 2018-04-18, 2018.

[439] D. D. Nguyen, L. Nguyen, and V. Sila, "Does Corporate Culture Affect Bank Risk-Taking? Evidence from Loan-Level Data", *British Journal of Management*, vol. 30, no. 1, pp. 106–133, 2019.

[440] F. Å. Nielsen, "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", *arXiv preprint arXiv:1103.2903*, 2011.

[441] NIH, *What is precision medicine?*, https://ghr.nlm.nih.gov/primer/precisionmedicine/definition, Accessed: 2020-08-07, 2020.

[442] R. A. Noe, J. R. Hollenbeck, B. Gerhart, and P. M. Wright, *Human resource management: Gaining a competitive advantage*. McGraw-Hill Education New York, NY, 2017.

[443] D. W. Nordstokke and B. D. Zumbo, "A new nonparametric Levene test for equal variances.", *Psicologica*, 2010.

[444] C. O'Reilly, "Corporations, culture, and commitment: Motivation and social control in organizations", *California management review*, vol. 31, no. 4, pp. 9–25, 1989.

[445] C. A. O'Reilly III, J. Chatman, and D. F. Caldwell, "People and organizational culture: A profile comparison approach to assessing person-organization fit", *Academy of management journal*, vol. 34, no. 3, pp. 487–516, 1991.

[446] R. M. Obérien, "A caution regarding rules of thumb for variance inflation factors", *Quality & quantity*, vol. 41, no. 5, pp. 673–690, 2007.

[447] H. J. Oh, C. Lauckner, J. Boehmer, R. Fewins-Bliss, and K. Li, "Facebooking for health: An examination into the solicitation and effects of health-related social support on social networking sites", *Computers in Human Behavior*, vol. 29, no. 5, pp. 2072–2080, 2013.

[448] R. Oldenburg, *The great good place: Cafes, coffee shops, bookstores, bars, hair salons, and other hangouts at the heart of a community*. Da Capo Press, 1999.

[449] C. L. Olson, "Practical considerations in choosing a manova test statistic: A rejoinder to stevens.", 1979.

[450] R. Olson, J. Verley, L. Santos, and C. Salas, "What we teach students about the Hawthorne studies: A review of content within a sample of introductory IO and OB textbooks", *The Industrial-Organizational Psychologist*, vol. 41, no. 3, pp. 23–39, 2004.

[451] A. Olteanu, C. Castillo, F. Diaz, and E. Kiciman, "Social data: Biases, methodological pitfalls, and ethical boundaries", 2016.

[452] ——, "Social data: Biases, methodological pitfalls, and ethical boundaries", *Frontiers in Big Data*, vol. 2, p. 13, 2019.

[453] E. Y. Ong, R. P. Ang, J. C. Ho, J. C. Lim, D. H. Goh, C. S. Lee, and A. Y. Chua, "Narcissism, extraversion and adolescents' self-presentation on facebook", *Personality and individual differences*, vol. 50, no. 2, pp. 180–185, 2011.

[454] J.-P. Onnela and S. L. Rauch, "Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health", *Neuropsychopharmacology*, vol. 41, no. 7, p. 1691, 2016.

[455] D. W. Organ, *Organizational citizenship behavior: The good soldier syndrome.* Lexington Books/DC Heath and Com, 1988.

[456] J. Ormel, A. Bastiaansen, H. Riese, E. H. Bos, M. Servaas, M. Ellenbogen, J. G. Rosmalen, and A. Aleman, "The biological and psychological basis of neuroticism: current status and future directions", *Neuroscience & Biobehavioral Reviews*, vol. 37, no. 1, pp. 59–72, 2013.

[457] A. Osman, J. L. Wong, C. L. Bagge, S. Freedenthal, P. M. Gutierrez, and G. Lozano, "The depression anxiety stress Scales—21 (DASS-21): further examination of dimensions, scale reliability, and correlates", *Journal of clinical psychology*, vol. 68, no. 12, pp. 1322–1338, 2012.

[458] D. Oswald, F. Sherratt, and S. Smith, "Handling the Hawthorne effect: The challenges surrounding a participant observer", *Review of social studies*, vol. 1, no. 1, pp. 53–73, 2014.

[459] X. Page, P. Wisniewski, B. P. Knijnenburg, and M. Namara, "Social media's have-nots: An era of social disenfranchisement", *Internet Research*, 2018.

[460] L. Palen, "Online social media in crisis events", *Educause Quarterly*, vol. 31, no. 3, pp. 76–78, 2008.

[461] H. J. Pandit and D. Lewis, "Ease and ethics of user profiling in black mirror", in *Companion Proceedings of the The Web Conference 2018*, 2018.

[462] G. Panger, "Social comparison in social media: A look at facebook and twitter", in *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, 2014, pp. 2095–2100.

[463] U. N. Pareek, *Training instruments for human resource development*. Tata McGraw-Hill, 1997.

[464] S. Pasch, "Corporate Culture and Industry-Fit: A Text Mining Approach", 2018.

[465] R. Patel, T. Lloyd, R. Jackson, M. Ball, H. Shetty, M. Broadbent, J. R. Geddes, R. Stewart, P. McGuire, and M. Taylor, "Mood instability is a common feature of mental health disorders and is associated with poor clinical outcomes", *BMJ open*, vol. 5, no. 5, e007504, 2015.

[466] T. Patel, C. R. Brewin, J. Wheatley, A. Wells, P. Fisher, and S. Myers, "Intrusive images and memories in major depression", *Behaviour research and therapy*, vol. 45, no. 11, pp. 2573–2580, 2007.

[467] E. H. Patzelt, C. A. Hartley, and S. J. Gershman, "Computational phenotyping: using models to understand individual differences in personality, development, and mental illness", *Personality Neuroscience*, vol. 1, 2018.

[468] M. J. Paul and M. Dredze, "You Are What You Tweet: Analyzing Twitter for Public Health.", in *ICWSM*, 2011.

[469] E. Pavlick and C. Callison-Burch, "The gun violence database", *arXiv preprint arXiv:1610.01670*, 2016.

[470] J. Payne, A. M. Chambers, and E. A. Kensinger, "Sleep promotes lasting changes in selective memory for emotional scenes", *Frontiers in integrative neuroscience*, vol. 6, p. 108, 2012.

[471] J. L. Pearce, "Bringing some clarity to role ambiguity research", *Academy of Management Review*, 1981.

[472] L. Peng, L. Chen, Z. Ye, and Y. Zhang, "AROMA: A Deep Multi-Task Learning Based Simple and Complex Human Activity Recognition Method Using Wearable Sensors", *PACM IMWUT*, 2018.

[473] J. W. Pennebaker and C. K. Chung, "Expressive writing, emotional upheavals, and health", *Handbook of health psychology*, pp. 263–284, 2007.

[474] J. W. Pennebaker, C. K. Chung, J. Frazee, G. M. Lavergne, and D. I. Beaver, "When small words foretell academic success: The case of college admissions essays", *PloS one*, vol. 9, no. 12, e115844, 2014.

[475] J. W. Pennebaker, M. R. Mehl, and K. G. Niederhoffer, "Psychological aspects of natural language use: Our words, our selves", *Annual review of psychology*, vol. 54, no. 1, pp. 547–577, 2003.

[476] J. Pennington, R. Socher, and C. Manning, “Glove: Global vectors for word representation”, in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.

[477] M. Petticrew, S. Cummins, C. Ferrell, A. Findlay, C. Higgins, C. Hoy, A. Kearns, and L. Sparks, “Natural experiments: an underused tool for public health?”, *Public health*, vol. 119, no. 9, pp. 751–757, 2005.

[478] Pew, *pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018*, 2019.

[479] ——, *Social Media Fact Sheet*, pewinternet.org/fact-sheet/social-media, 2019.

[480] S. Picazo-Vela, S. Y. Chou, A. J. Melcher, and J. M. Pearson, “Why provide an online review? an extended theory of planned behavior and the role of big-five personality traits”, *Computers in Human Behavior*, vol. 26, no. 4, pp. 685–696, 2010.

[481] J. Pierce, “Undesigning technology: Considering the negation of design by design”, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012, pp. 957–966.

[482] E. Pitler and A. Nenkova, “Revisiting readability: A unified framework for predicting text quality”, in *Proceedings of the conference on empirical methods in natural language processing*, Association for Computational Linguistics, 2008, pp. 186–195.

[483] S. Poijula, K.-E. Wahlberg, and A. Dyregrov, “Adolescent suicide and suicide contagion in three secondary schools”, *International Journal of Emergency Mental Health*, vol. 3, no. 3, pp. 163–170, 2001.

[484] M. Poland and D. C. Mare, “Defining geographic communities”, 2005.

[485] J. P. Pollak, P. Adams, and G. Gay, “PAM: a photographic affect meter for frequent, in situ measurement of affect”, in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 2011, pp. 725–734.

[486] D. E. Poplin, *Communities: A survey of theories and methods of research*. Macmillan, 1979.

[487] J. Posner, J. A. Russell, and B. S. Peterson, “The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology”, *Development and psychopathology*, vol. 17, no. 3, pp. 715–734, 2005.

[488] M. I. Posner and M. K. Rothbart, "Developing mechanisms of self-regulation", *Development and psychopathology*, vol. 12, no. 03, pp. 427–441, 2000.

[489] D. I. Prajogo and C. M. McDermott, "The relationship between multidimensional organizational culture and performance", *International Journal of Operations & Production Management*, vol. 31, no. 7, pp. 712–735, 2011.

[490] J. O. Prochaska, S. Johnson, and P. Lee, "The transtheoretical model of behavior change.", 1998.

[491] D. Prothrow-Stith and S. Quaday, "Hidden Casualties: The Relationship between Violence and Learning.", in *Streamlined Seminar*, ERIC, vol. 14, 1995, n2.

[492] Y. Qu, C. Huang, P. Zhang, and J. Zhang, "Microblogging after a major disaster in China: a case study of the 2010 Yushu earthquake", in *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, ACM, 2011, pp. 25–34.

[493] T. Quatman and C. Swanson, "Academic self-disclosure in adolescence", *Genetic, social, and general psychology monographs*, vol. 128, no. 1, pp. 47–76, 2002.

[494] D. Quercia, M. Kosinski, D. Stillwell, and J. Crowcroft, "Our twitter profiles, our selves: Predicting personality with twitter", in *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*, IEEE, 2011, pp. 180–185.

[495] R. E. Quinn and J. Rohrbaugh, "A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis", *Management science*, vol. 29, no. 3, pp. 363–377, 1983.

[496] M. A. Quińones, J. K. Ford, and M. S. Teachout, "The relationship between work experience and job performance: A conceptual and meta-analytic review", *Personnel psychology*, vol. 48, no. 4, pp. 887–910, 1995.

[497] V. Radu, C. Tong, S. Bhattacharya, N. D. Lane, C. Mascolo, M. K. Marina, and F. Kawsar, "Multimodal deep learning for activity and context recognition", *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, pp. 1–27, 2018.

[498] M. Raghavan, S. Barocas, J. Kleinberg, and K. Levy, "Mitigating bias in algorithmic hiring: Evaluating claims and practices", in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pp. 469–481.

[499] T. Rahman, A. T. Adams, M. Zhang, E. Cherry, B. Zhou, H. Peng, and T. Choudhury, "BodyBeat: a mobile system for sensing non-speech body sounds.", in *MobiSys*, 2014.

[500] L. Razmerita, K. Kirchner, and P. Nielsen, “What factors influence knowledge sharing in organizations? A social dilemma perspective of social media communication”, *Journal of knowledge Management*, 2016.

[501] L. Reinecke and S. Trepte, “Authenticity and well-being on social network sites: A two-wave longitudinal study on the effects of online authenticity and the positivity bias in SNS communication”, *Computers in Human Behavior*, vol. 30, pp. 95–102, 2014.

[502] N. Rekabsaz, M. Lupu, and A. Hanbury, “Exploration of a threshold for similarity based on uncertainty in word embedding”, in *European Conference on Information Retrieval*, Springer, 2017, pp. 396–409.

[503] P. Release, *Census Bureau Reports*, https://www.census.gov/newsroom/press-releases/2016/cb16-139.html, Accessed: 2020-04-04, 2016.

[504] P. Resnik, W. Armstrong, L. Claudino, T. Nguyen, V.-A. Nguyen, and J. Boyd-Graber, “Beyond lda: Exploring supervised topic modeling for depression-related language in twitter”, in *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, 2015, pp. 99–107.

[505] R. E. Rhodes and N. Smith, “Personality correlates of physical activity: a review and meta-analysis”, *British journal of sports medicine*, vol. 40, no. 12, pp. 958–965, 2006.

[506] M. Richey, A. Gonibeed, and M. Ravishankar, “The perils and promises of self-disclosure on social media”, *Information Systems Frontiers*, vol. 20, no. 3, pp. 425–437, 2018.

[507] H. R. Riggio and R. E. Riggio, “Emotional expressiveness, extraversion, and neuroticism: A meta-analysis”, *Journal of Nonverbal Behavior*, vol. 26, no. 4, pp. 195–218, 2002.

[508] K.-J. Ritter, R. A. Matthews, M. T. Ford, and A. A. Henderson, “Understanding role stressors and job satisfaction over time using adaptation theory.”, *Journal of Applied Psychology*, vol. 101, no. 12, p. 1655, 2016.

[509] J. R. Rizzo, R. J. House, and S. I. Lirtzman, “Role conflict and ambiguity in complex organizations”, *Administrative science quarterly*, pp. 150–163, 1970.

[510] W. Roberts and J. Strayer, “Empathy, emotional expressiveness, and prosocial behavior”, *Child development*, vol. 67, no. 2, pp. 449–470, 1996.

[511] P. Robles-Granda, S. Lin, X. Wu, S. D'Mello, G. J. Martinez, K. Saha, K. Nies, G. Mark, A. T. Campbell, M. De Choudhury, *et al.*, "Jointly Predicting Job Performance, Personality, Cognitive Ability, Affect, and Well-Being", *arXiv preprint arXiv:2006.08364*, 2020.

[512] J. E. Roeckelein, *Dictionary of theories, laws, and concepts in psychology*. Greenwood Publishing Group, 1998.

[513] F. J. Roethlisberger and W. J. Dickson, *Management and the Worker*. Psychology press, 2003, vol. 5.

[514] J. Rooksby, A. Morrison, and D. Murray-Rust, "Student perspectives on digital phenotyping: The acceptability of using smartphone data to assess mental health", in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–14.

[515] P. R. Rosenbaum and D. B. Rubin, "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score", *The American Statistician*, vol. 39, no. 1, pp. 33–38, 1985.

[516] I. M. Rosenstock, "Historical origins of the health belief model", *Health education monographs*, vol. 2, no. 4, pp. 328–335, 1974.

[517] I. M. Rosenstock, V. J. Strecher, and M. H. Becker, "Social learning theory and the health belief model", *Health education quarterly*, vol. 15, no. 2, pp. 175–183, 1988.

[518] S. E. Ross, B. C. Niebling, and T. M. Heckert, "Sources of stress among college students", *Social psychology*, vol. 61, no. 5, pp. 841–846, 1999.

[519] M. Rotundo and P. R. Sackett, "The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy-capturing approach.", *Journal of applied psychology*, vol. 87, no. 1, p. 66, 2002.

[520] D. Rousseau, "Quantitative assessment of organizational culture: The case for multiple", 1990.

[521] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis", *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.

[522] D. W. Ruck, S. K. Rogers, M. Kabrisky, M. E. Oxley, and B. W. Suter, "The multilayer perceptron as an approximation to a Bayes optimal discriminant function", *IEEE Trans. Neural Netw.*,

[523] S. Rude, E.-M. Gortner, and J. Pennebaker, "Language use of depressed and depression-vulnerable college students", *Cognition & Emotion*, vol. 18, no. 8, pp. 1121–1133, 2004.

[524] O. Rudovic, J. Lee, M. Dai, B. Schuller, and R. W. Picard, "Personalized machine learning for robot perception of affect and engagement in autism therapy", *Science Robotics*, vol. 3, no. 19, 2018.

[525] D. Ruths and J. Pfeffer, "Social media for large studies of behavior", *Science*, vol. 346, no. 6213, pp. 1063–1064, 2014.

[526] A. Sadilek, H. Kautz, and V. Silenzio, "Modeling spread of disease from social interactions", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 6, 2012.

[527] R. Safdari, E. Maserat, H. A. Aghdaei, *et al.*, "Person centered prediction of survival in population based screening program by an intelligent clinical decision support system", *Gastroenterology and Hepatology from bed to Bench*, 2017.

[528] K. Saha, A. E. Bayraktaroglu, A. T. Campbell, N. V. Chawla, M. De Choudhury, S. K. D'Mello, A. K. Dey, G. Gao, J. M. Gregg, K. Jagannath, G. Mark, G. J. Martinez, S. M. Mattingly, E. Moskal, A. Sirigiri, A. Striegel, and D. W. Yoo, "Social media as a passive sensor in longitudinal studies of human behavior and wellbeing", in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–8.

[529] K. Saha, L. Chan, K. De Barbaro, G. D. Abowd, and M. De Choudhury, "Inferring mood instability on social media by leveraging ecological momentary assessments", *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, p. 95, 2017.

[530] K. Saha, E. Chandrasekharan, and M. De Choudhury, "Prevalence and psychological effects of hateful speech in online college communities", in *Proceedings of the 10th ACM Conference on Web Science*, 2019, pp. 255–264.

[531] K. Saha and M. De Choudhury, "Modeling stress with social media around incidents of gun violence on college campuses", *PACM Human-Computer Interaction*, no. CSCW, 2017.

[532] K. Saha, T. Grover, S. M. Mattingly, V. Das Swain, P. Gupta, G. J. Martinez, P. Robles-Granda, G. Mark, A. Striegel, and M. De Choudhury, "Person-centered predictions of psychological constructs with social media contextualized by multimodal sensing", *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 1, pp. 1–32, 2021.

[533] K. Saha, S. C. Kim, M. D. Reddy, A. J. Carter, E. Sharma, O. L. Haimson, and M. De Choudhury, "The language of LGBTQ+ minority stress experiences on social media", *PACM HCI*, no. CSCW, 2019.

[534] K. Saha, Y. Liu, N. Vincent, F. A. Chowdhury, L. Neves, N. Shah, and M. W. Bos, "Advertiming matters: Examining user ad consumption for effective ad allocations on social media", in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–18.

[535] K. Saha, M. D. Reddy, and M. De Choudhury, "JobLex: A Lexico-Semantic Knowledgebase of Occupational Information Descriptors", in *SocInfo*, 2019.

[536] K. Saha, M. D. Reddy, S. Mattingly, E. Moskal, A. Sirigiri, and M. De Choudhury, "Libra: On linkedin based role ambiguity and its relationship with wellbeing and job performance", *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. CSCW, pp. 1–30, 2019.

[537] K. Saha, M. D. Reddy, V. Das Swain, J. M. Gregg, T. Grover, S. Lin, G. J. Martinez, S. M. Mattingly, S. Mirjafari, R. Mulukutla, K. Nies, P. Robles-Granda, A. Sirigiri, D. W. Yoo, P. Audia, A. T. Campbell, N. V. Chawla, S. K. D'Mello, A. K. Dey, K. Jiang, Q. Liu, G. Mark, E. Moskal, A. Striegel, and M. De Choudhury, "Imputing missing social media data stream in multisensor studies of human behavior", in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, IEEE, 2019, pp. 178–184.

[538] K. Saha, J. Seybolt, S. M. Mattingly, T. Aledavood, C. Konjeti, G. J. Martinez, T. Grover, G. Mark, and M. De Choudhury, "What life events are disclosed on social media, how, when, and by whom?", in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–22.

[539] K. Saha and A. Sharma, "Causal factors of effective psychosocial outcomes in online mental health communities", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 14, 2020, pp. 590–601.

[540] K. Saha, B. Sugar, J. Torous, B. Abrahao, E. Kıcıman, and M. De Choudhury, "A social media study on the effects of psychiatric medication use", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 13, 2019, pp. 440–451.

[541] K. Saha, J. Torous, E. D. Caine, and M. De Choudhury, "Psychosocial effects of the covid-19 pandemic: Large-scale quasi-experimental study on social media", *Journal of medical internet research*, vol. 22, no. 11, e22600, 2020.

[542] K. Saha, J. Torous, S. K. Ernala, C. Rizuto, A. Stafford, and M. De Choudhury, "A computational study of mental health awareness campaigns on social media", *Translational behavioral medicine*, vol. 9, no. 6, pp. 1197–1207, 2019.

[543] K. Saha, I. Weber, M. L. Birnbaum, and M. De Choudhury, "Characterizing awareness of schizophrenia among facebook users by leveraging facebook advertisement estimates", *Journal of Medical Internet Research*, vol. 19, no. 5, e156, 2017.

[544] K. Saha, I. Weber, and M. De Choudhury, "A social media based examination of the effects of counseling recommendations after student deaths on college campuses", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 12, 2018.

[545] K. Saha, A. Yousuf, R. Boyd, J. W. Pennebaker, and M. De Choudhury, "Mental health consultations on college campuses: Examining the predictive ability of social media", *Available at SSRN 3774189*, 2021.

[546] K. Saha, A. Yousuf, L. Hickman, P. Gupta, L. Tay, and M. De Choudhury, "A social media study on demographic differences in perceived job satisfaction", *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–29, 2021.

[547] J. F. Salgado, "The five factor model of personality and job performance in the European community.", *Journal of Applied psychology*, vol. 82, no. 1, p. 30, 1997.

[548] J. F. Sallis and N. Owen, *Physical activity and behavioral medicine*. SAGE publications, 1998, vol. 3.

[549] W. R. Saltzman, R. S. Pynoos, C. M. Layne, A. M. Steinberg, and E. Aisenberg, "Trauma-and grief-focused intervention for adolescents exposed to community violence: Results of a school-based screening and group treatment protocol.", *Group Dynamics: Theory, Research, and Practice*, vol. 5, no. 4, p. 291, 2001.

[550] E. Sandvik, E. Diener, and L. Seidlitz, "Subjective well-being: The convergence and stability of self-report and non-self-report measures", in *Assessing well-being*, Springer, 2009, pp. 119–138.

[551] M. Sapone, *Why titles will kill your startup*, medium.com/@MsSapone/why-titles-will-kill-your-startup-58402c5b6954, Accessed: 2019-03-26, 2019.

[552] V. Satopaa, J. Albrecht, D. Irwin, and B. Raghavan, "Finding a" kneedle" in a haystack: Detecting knee points in system behavior", in *ICDCS*, 2011.

[553] M. Z. Satti, T. M. Khan, *et al.*, "Association of physical activity and sleep quality with academic performance among fourth-year MBBS students of Rawalpindi Medical University", *Cureus*, vol. 11, no. 7, 2019.

[554] W. B. Schaufeli and M. C. Peeters, "Job stress and burnout among correctional officers: A literature review", *International Journal of Stress Management*, vol. 7, no. 1, pp. 19–48, 2000.

[555] D. A. Scheufele and D. Tewksbury, "Framing, agenda setting, and priming: The evolution of three media effects models", *Journal of communication*, vol. 57, no. 1, pp. 9–20, 2007.

[556] S. Schipolowski, O. Wilhelm, and U. Schroeders, "On the nature of crystallized intelligence: The relationship between verbal ability and factual knowledge", *Intelligence*, vol. 46, pp. 156–168, 2014.

[557] F. L. Schmidt, J. E. Hunter, and A. N. Outerbridge, "Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance.", *Journal of applied psychology*, vol. 71, no. 3, p. 432, 1986.

[558] S. Schmidt, U. Roesler, T. Kusserow, and R. Rau, "Uncertainty in the workplace: examining role ambiguity and role conflict, and their link to depression – a meta-analysis", *European Journal of Work and Organizational Psychology*, vol. 23, no. 1, pp. 91–106, 2014.

[559] W. J. Schneider and K. S. McGrew, "The Cattell-Horn-Carroll model of intelligence.", 2012.

[560] N. Schneiderman, G. Ironson, and S. D. Siegel, "Stress and health: psychological, behavioral, and biological determinants", *Annu. Rev. Clin. Psychol.*, vol. 1, pp. 607–628, 2005.

[561] J. Schrammel, C. Köffel, and M. Tscheligi, "Personality traits, usage patterns and information disclosure in online communities", *People and Computers XXIII Celebrating People and Technology*, pp. 169–174, 2009.

[562] H. Schut Margaret Stroebe, "The dual process model of coping with bereavement: Rationale and description", *Death studies*, vol. 23, no. 3, pp. 197–224, 1999.

[563] A. J. Schwartz and L. C. Whitaker, "Suicide among college students: Assessment, treatment, and intervention.", 1990.

[564] H. A. Schwartz, J. C. Eichstaedt, M. L. Kern, *et al.*, "Personality, gender, and age in the language of social media: The open-vocabulary approach", *PloS one*, vol. 8, no. 9, e73791, 2013.

[565] C. N. Scollon, C.-K. Prieto, and E. Diener, "Experience sampling: promises and pitfalls, strength and weaknesses", in *Assessing well-being*, Springer, 2009, pp. 157–180.

[566] G. A. Seber and A. J. Lee, *Linear regression analysis*. John Wiley & Sons, 2012, vol. 329.

[567] T. E. Seeman, “Social ties and health: The benefits of social integration”, *Annals of epidemiology*, vol. 6, no. 5, pp. 442–451, 1996.

[568] E. A. Seemann, W. C. Buboltz, A. Thomas, B. Soper, and L. Wilkinson, “Normal personality variables and their relationship to psychological reactance.”, *Individual Differences Research*, vol. 3, no. 2, 2005.

[569] G. Seidman, “Self-presentation and belonging on facebook: How personality influences social media use and motivations”, *Personality and individual differences*, vol. 54, no. 3, pp. 402–407, 2013.

[570] H. Selye, “The stress of life.”, 1956.

[571] J. A. Shaffer, A. Li, and J. Bagger, “A moderated mediation model of personality, self-monitoring, and ocb”, *Human Performance*, vol. 28, no. 2, pp. 93–111, 2015.

[572] N. S. Shami, M. Muller, A. Pal, M. Masli, and W. Geyer, “Inferring employee engagement from social media”, in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 2015, pp. 3999–4008.

[573] N. S. Shami, J. Nichols, and J. Chen, “Social media participation and performance at work: a longitudinal study”, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2014, pp. 115–118.

[574] N. S. Shami, J. Yang, L. Panc, C. Dugan, T. Ratchford, J. C. Rasmussen, Y. M. Assogba, T. Steier, T. Soule, S. Lupushor, *et al.*, “Understanding employee social media chatter with enterprise social pulse”, in *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, ACM, 2014, pp. 379–392.

[575] P. Sheldon, “Examining gender differences in self-disclosure on facebook versus face-to-face”, *The Journal of Social Media in Society*, vol. 2, no. 1, 2013.

[576] A. Shepitsen, J. Gemmell, B. Mobasher, and R. Burke, “Personalized recommendation in social tagging systems using hierarchical clustering”, in *Proceedings of the 2008 ACM conference on Recommender systems*, 2008, pp. 259–266.

[577] S. Shiffman, A. A. Stone, and M. R. Hufford, “Ecological momentary assessment”, *Annu. Rev. Clin. Psychol.*, vol. 4, pp. 1–32, 2008.

[578] W. C. Shipley, *Shipley-2: manual*. WPS, 2009.

[579] E. Shmueli, V. K. Singh, B. Lepri, and A. Pentland, "Sensing, understanding, and shaping social behavior", *IEEE Transactions on Computational Social Systems*, vol. 1, no. 1, pp. 22–34, 2014.

[580] J. Silvester, N. R. Anderson, and F. Patterson, "Organizational culture change: An inter-group attributional analysis", *Journal of Occupational and Organizational Psychology*, vol. 72, no. 1, pp. 1–23, 1999.

[581] J. Singh, "Facets , Ambiguity :" *The Journal of Marketing*, vol. 57, no. 2, pp. 11–31, 1993.

[582] M. M. Skeels and J. Grudin, "When social networks cross boundaries: a case study of workplace use of facebook and linkedin", in *Proceedings of the ACM 2009 international conference on Supporting group work*, ACM, 2009, pp. 95–104.

[583] C. S. Smith, J. Tisak, and R. A. Schmieder, "The measurement properties of the role conflict and role ambiguity scales: A review and extension of the empirical research", *Journal of organizational Behavior*, 1993.

[584] I. D. Smith, S. J. Coombs, *et al.*, "The Hawthorne Effect: is it a help or a hindrance in social science research?", *Change (Sydney, NSW)*, vol. 6, no. 1, p. 97, 2003.

[585] J. Snoek, H. Larochelle, and R. P. Adams, "Practical bayesian optimization of machine learning algorithms", in *Advances in neural information processing systems*, 2012, pp. 2951–2959.

[586] Snoopsnoo, *Reddit User and Subreddit Analytics: snoopsnoo.com*, Accessed: 2018-04-18, 2018.

[587] M. Snyder, "Self-monitoring processes", in *Advances in experimental social psychology*, vol. 12, Elsevier, 1979, pp. 85–128.

[588] D. J. Solove, "I've got nothing to hide and other misunderstandings of privacy", *San Diego L. Rev.*, 2007.

[589] C. J. Soto and O. P. John, "The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power.", *Journal of Personality and Social Psychology*, vol. 113, no. 1, p. 117, 2017.

[590] K. Sparks, C. Cooper, Y. Fried, and A. Shirom, "The effects of hours of work on health: a meta-analytic review", *Journal of occupational and organizational psychology*, vol. 70, no. 4, pp. 391–408, 1997.

[591] C. D. Spielberger, F. Gonzalez-Reigosa, A. Martinez-Urrutia, L. F. Natalicio, and D. S. Natalicio, "The state-trait anxiety inventory", *Revista Interamerican Journal of Psychology*, 2017.

[592] E. S. Spiro, "Research opportunities at the intersection of social media and survey data", *Current Opinion in Psychology*, vol. 9, pp. 67–71, 2016.

[593] D. Spruijt-Metz and W. Nilsen, "Dynamic models of behavior for just-in-time adaptive interventions", *IEEE Pervasive Computing*, no. 3, pp. 13–17, 2014.

[594] A. Srinivasan, H. Guo, and S. Devaraj, "Who cares about your big day? impact of life events on dynamics of social networks", *Decision Sciences*, vol. 48, no. 6, pp. 1062–1097, 2017.

[595] C. Stachl, Q. Au, R. Schoedel, S. D. Gosling, G. M. Harari, D. Buschek, S. T. Völkel, T. Schuwerk, M. Oldemeier, T. Ullmann, *et al.*, "Predicting personality from patterns of behavior collected with smartphones", *Proceedings of the National Academy of Sciences*, vol. 117, no. 30, pp. 17 680–17 687, 2020.

[596] D. R. Stalder, "Need for closure, the big five, and public self-consciousness", *The Journal of social psychology*, vol. 147, no. 1, pp. 91–94, 2007.

[597] K. Starbird, L. Palen, A. L. Hughes, and S. Vieweg, "Chatter on the red: What hazards threat reveals about the social life of microblogged information", in *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, 2010, pp. 241–250.

[598] C. Steindl, E. Jonas, S. Sittenthaler, E. Traut-Mattausch, and J. Greenberg, "Understanding psychological reactance", *Zeitschrift für Psychologie*, 2015.

[599] C. Steinfield, N. B. Ellison, and C. Lampe, "Social capital, self-esteem, and use of online social network sites: A longitudinal analysis", *JADP*, 2008.

[600] J. H. Steward, *Theory of culture change: The methodology of multilinear evolution*. University of Illinois Press, 1972.

[601] R. Strack, C. Von Der Linden, M. Booker, and A. Strohmayr, "Decoding global talent", *BCG Perspectives*, 2014.

[602] A. Strauss and J. Corbin, "Open coding", *Basics of qualitative research: Grounded theory procedures and techniques*, 1990.

[603] H. Suh, N. Shahriaree, E. B. Hekler, and J. A. Kientz, "Developing and Validating the User Burden Scale: A Tool for Assessing User Burden in Computing Systems", in

Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, ACM, 2016, pp. 3988–3999.

[604] *Suicide now second-leading cause of death among college students*, tinyurl.com/yb4hhus4, Accessed: 2017-10-26, 2017.

[605] S. E. Sullivan and R. S. Bhagat, "Organizational stress, job satisfaction and job performance: where do we go from here?", *Journal of management*, vol. 18, no. 2, pp. 353–374, 1992.

[606] R. Sumbaly, J. Kreps, and S. Shah, "The big data ecosystem at linkedin", in *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*, 2013, pp. 1125–1134.

[607] K. Suresh and S. Chandrashekara, "Sample size estimation and power analysis for clinical research studies", *Journal of human reproductive sciences*, vol. 5, no. 1, p. 7, 2012.

[608] J. Swan and P. M. Hamilton, "Mental Health Crises", 2017.

[609] "Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects", *Journal of clinical epidemiology*, vol. 67, no. 3, pp. 267–277, 2014.

[610] P. Tambe and L. M. Hitt, "Now IT's personal: Offshoring and the shifting skill composition of the US information technology workforce", *Management Science*, vol. 58, no. 4, pp. 678–695, 2012.

[611] B. Tancer, *Everyone's a critic: Winning customers in a review-driven world*. Penguin, 2014.

[612] T. L.-P. Tang and P. R. Gilbert, "Attitudes toward money as related to intrinsic and extrinsic job satisfaction, stress and work-related attitudes", *Personality and Individual Differences*, vol. 19, no. 3, pp. 327–332, 1995.

[613] Y. R. Tausczik and J. W. Pennebaker, "The psychological meaning of words: LIWC and computerized text analysis methods", *Journal of language and social psychology*, vol. 29, no. 1, pp. 24–54, 2010.

[614] R. P. Tett, D. N. Jackson, and M. Rothstein, "Personality measures as predictors of job performance: A meta-analytic review", *Personnel psychology*, vol. 44, no. 4, pp. 703–742, 1991.

[615] J. Thom-Santelli, D. R. Millen, and D. Gergle, “Organizational acculturation and social networking”, in *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, ACM, 2011, pp. 313–316.

[616] S. Thomée, A. Härenstam, and M. Hagberg, “Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults-a prospective cohort study”, *BMC public health*, vol. 11, no. 1, p. 66, 2011.

[617] E. R. Thompson, “Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS)”, *Journal of cross-cultural psychology*, vol. 38, no. 2, pp. 227–242, 2007.

[618] P. W. Thorndyke, “Cognitive structures in comprehension and memory of narrative discourse”, *Cognitive psychology*, vol. 9, no. 1, pp. 77–110, 1977.

[619] S. Tifferet and I. Vilnai-Yavetz, “Self-presentation in LinkedIn portraits: common features, gender, and occupational differences”, *Computers in Human Behavior*, vol. 80, pp. 33–48, 2018.

[620] R. Tourangeau, L. J. Rips, and K. Rasinski, *The psychology of survey response*. 2000.

[621] H. C. Triandis, “Subjective culture and interpersonal relations across cultures”, *Annals of the New York Academy of Sciences*, vol. 285, no. 1, pp. 418–434, 1977.

[622] T. J. Trull, M. B. Solhan, S. L. Tragesser, S. Jahng, P. K. Wood, T. M. Piasecki, and D. Watson, “Affective instability: measuring a core feature of borderline personality disorder with ecological momentary assessment.”, *Journal of abnormal psychology*, vol. 117, no. 3, p. 647, 2008.

[623] Z. Tufekci, “Can you see me now? audience and disclosure regulation in online social network sites”, *Bulletin of Science, Technology & Society*, vol. 28, no. 1, pp. 20–36, 2008.

[624] ——, “Big questions for social media big data: Representativeness, validity and other methodological pitfalls”, in *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.

[625] J. C. Turner, E. V. Leno, and A. Keller, “Causes of mortality among American college students: a pilot study”, *Journal of college student psychotherapy*, vol. 27, no. 1, pp. 31–42, 2013.

[626] J. M. Twenge and B. H. Spitzberg, “Declines in non-digital social interaction among americans, 2003–2017”, *Journal of Applied Social Psychology*, 2020.

[627] J. Tyler, "The Myth of the Ninja Rockstar Developer", in *Building Great Software Engineering Teams*, 2015.

[628] S. Utz, "Is LinkedIn making you more successful? The informational benefits derived from public social media", *New Media & Society*, vol. 18, no. 11, pp. 2685–2702, 2016.

[629] S. Utz and J. Breuer, "The Relationship Between Networking, LinkedIn Use, and Retrieving Informational Benefits", *Cyberpsychology, Behavior, and Social Networking*, 2019.

[630] P. M. Valkenburg, J. Peter, and A. P. Schouten, "Friend networking sites and their relationship to adolescents' well-being and social self-esteem", *CyberPsychology & Behavior*, vol. 9, no. 5, pp. 584–590, 2006.

[631] T. F. Van de Mortel *et al.*, "Faking it: social desirability response bias in self-report research", *Australian Journal of Advanced Nursing, The*, vol. 25, no. 4, p. 40, 2008.

[632] M. Van der Laan, K. Pollard, and J. Bryan, "A new partitioning around medoids algorithm", *Journal of Statistical Computation and Simulation*, vol. 73, no. 8, pp. 575–584, 2003.

[633] J. Van Dijck, "'You have one identity': Performing the self on Facebook and LinkedIn", *Media, culture & society*, vol. 35, no. 2, pp. 199–215, 2013.

[634] L. Van Dyne, E. Kossek, and S. Lobel, "Less need to be there: Cross-level effects of work practices that support work-life flexibility and enhance group processes and group-level OCB", *Human Relations*, 2007.

[635] L. Van Dyne, D. Vandewalle, T. Kostova, M. E. Latham, and L. Cummings, "Collectivism, propensity to trust and self-esteem as predictors of organizational citizenship in a non-work setting", *Journal of organizational behavior*, vol. 21, no. 1, pp. 3–23, 2000.

[636] O. Van Reeth, L. Weibel, K. Spiegel, R. Leproult, C. Dugovic, and S. Maccari, "Physiology of sleep (review)–interactions between stress and sleep: from basic research to clinical situations", *Sleep Medicine Reviews*, vol. 4, no. 2, pp. 201–219, 2000.

[637] H. Van Remoortel, C. A. Camillo, D. Langer, M. Hornikx, H. Demeyer, C. Burtin, M. Decramer, R. Gosselink, W. Janssens, and T. Troosters, "Moderate intense physical activity depends on selected Metabolic Equivalent of Task (MET) cut-off and type of data analysis", *PLoS One*, vol. 8, no. 12, e84365, 2013.

[638] J. Van Scotter, S. J. Motowidlo, and T. C. Cross, “Effects of task performance and contextual performance on systemic rewards.”, *Journal of applied psychology*, vol. 85, no. 4, p. 526, 2000.

[639] M. Van Sell, A. P. Brief, and R. S. Schuler, “Role Conflict and Role Ambiguity: Integration of the Literature and Directions for Future Research”, *Human Relations*, vol. 34, no. 1, pp. 43–71, 1981.

[640] Y. Vardi, “The effects of organizational and ethical climates on misconduct at work”, *Journal of Business ethics*, vol. 29, no. 4, pp. 325–337, 2001.

[641] A. Vehtari, A. Gelman, and J. Gabry, “Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC”, *Statistics and Computing*, vol. 27, no. 5, pp. 1413–1432, 2017.

[642] N. van de Ven, A. Bogaert, A. Serlie, M. J. Brandt, and J. J. Denissen, “Personality perception based on LinkedIn profiles”, *Journal of Managerial Psychology*, vol. 32, no. 6, pp. 418–429, 2017.

[643] C. Viswesvaran and D. S. Ones, “Perspectives on models of job performance”, *International Journal of Selection and Assessment*, vol. 8, no. 4, pp. 216–226, 2000.

[644] J. Vitak and J. Kim, “‘‘you can’t block people offline”: Examining how facebook’s affordances shape the disclosure process”, in *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, 2014.

[645] J. Vitak, C. Lampe, R. Gray, and N. B. Ellison, “‘‘why won’t you be my facebook friend?” strategies for managing context collapse in the workplace”, in *Proceedings of the 2012 iConference*, 2012, pp. 555–557.

[646] M. Vlachos, P. Yu, and V. Castelli, “On periodicity detection and structural periodic similarity”, in *Proceedings of the 2005 SIAM International Conference on Data Mining*, SIAM, 2005, pp. 449–460.

[647] J. A. Walcott-McQuigg, J. J. Zerwic, A. Dan, and M. A. Kelley, “An ecological approach to physical activity in African American women.”, *Medscape women’s health*, vol. 6, no. 6, pp. 3–3, 2001.

[648] H. M. Wallach, I. Murray, R. Salakhutdinov, and D. Mimno, “Evaluation methods for topic models”, in *Proceedings of the 26th annual international conference on machine learning*, 2009, pp. 1105–1112.

[649] S. Wallot, “Recurrence quantification analysis of processes and products of discourse: A tutorial in R”, *Discourse Processes*, vol. 54, no. 5-6, pp. 382–405, 2017.

[650] R. Wang, M. S. Aung, S. Abdullah, R. Brian, A. T. Campbell, T. Choudhury, M. Hauser, J. Kane, M. Merrill, E. A. Scherer, *et al.*, "CrossCheck: Toward passive sensing and detection of mental health changes in people with schizophrenia", in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 2016, pp. 886–897.

[651] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones", in *Ubicomp*.

[652] R. Wang, G. Harari, P. Hao, X. Zhou, and A. T. Campbell, "SmartGPA: how smartphones can assess and predict academic performance of college students", in *Proc. Ubicomp*, 2015.

[653] W. Wang, G. M. Harari, R. Wang, S. R. Müller, S. Mirjafari, K. Masaba, and A. T. Campbell, "Sensing behavioral change over time: Using within-person variability features from mobile sensing to predict personality traits", *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 1–21, 2018.

[654] Y. Wang, G. Norcie, S. Komanduri, A. Acquisti, P. G. Leon, and L. F. Cranor, "‘‘i regretted the minute i pressed share" a qualitative study of regrets on facebook", in *Proceedings of the seventh symposium on usable privacy and security*, 2011, pp. 1–16.

[655] Y. Wang, M. Niiya, G. Mark, S. M. Reich, and M. Warschauer, "Coming of Age (Digitally): An Ecological View of Social Media Use among College Students", in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 2015.

[656] J. H. Ward Jr, "Hierarchical grouping to optimize an objective function", *Journal of the American statistical association*, vol. 58, no. 301, pp. 236–244, 1963.

[657] G. J. Warheit, "Life events, coping, stress, and depressive symptomatology", *American Journal of Psychiatry*, vol. 136, no. 4B, pp. 502–507, 1979.

[658] D. Watson and L. A. Clark, "The PANAS-X: Manual for the positive and negative affect schedule-expanded form", 1999.

[659] D. Watson, J. W. Pennebaker, and R. Folger, "Beyond negative affectivity: Measuring stress and satisfaction in the workplace", *Journal of Organizational Behavior Management*, vol. 8, no. 2, pp. 141–158, 1987.

[660] C. L. Webber Jr and J. P. Zbilut, “Recurrence quantification analysis of nonlinear dynamical systems”, *Tutorials in contemporary nonlinear methods for the behavioral sciences*, vol. 94, no. 2005, pp. 26–94, 2005.

[661] P. Welch, “The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms”, *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[662] A. Whiting and D. Williams, “Why people use social media: A uses and gratifications approach”, *Qualitative Market Research: An International Journal*, 2013.

[663] L. J. Williams and S. E. Anderson, “Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors”, *Journal of management*, vol. 17, no. 3, pp. 601–617, 1991.

[664] R. E. Wilson, S. D. Gosling, and L. T. Graham, “A review of facebook research in the social sciences”, *Perspectives on psychological science*, vol. 7, no. 3, pp. 203–220, 2012.

[665] K. Wise, B. Hamman, and K. Thorson, “Moderation, response rate, and message interactivity: Features of online communities and their effects on intent to participate”, *Journal of Computer-Mediated Communication*, vol. 12, no. 1, pp. 24–41, 2006.

[666] P. Wisniewski, H. Lipford, and D. Wilson, “Fighting for my space: Coping mechanisms for SNS boundary regulation”, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012, pp. 609–618.

[667] T.-T. Wong, “Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation”, *Pattern Recognition*, vol. 48, no. 9, pp. 2839–2846, 2015.

[668] J. Wood, D. W. Foy, C. Layne, R. Pynoos, and C. B. James, “An examination of the relationships between violence exposure, posttraumatic stress symptomatology, and delinquent activity: An “ecopathological” model of delinquent behavior among incarcerated adolescents”, *Journal of Aggression, Maltreatment & Trauma*, vol. 6, no. 1, pp. 127–147, 2002.

[669] R. Wrenn, “The grieving college student”, *Living with grief: At work, at school, at worship*, pp. 131–141, 1999.

[670] A. G. Wright and W. C. Woods, “Personalized models of psychopathology”, *Annual review of clinical psychology*, vol. 16, 2020.

[671] T. A. Wright and R. Cropanzano, "Psychological well-being and job satisfaction as predictors of job performance.", *Journal of occupational health psychology*, vol. 5, no. 1, p. 84, 2000.

[672] wsj, *Real People. Real Reviews. Real Extortion Scheme?, https://blogs.wsj.com/law/2010/02/26/real-people-real-reviews-real-extortion-scheme/*, Accessed: 2019-03-28, 2019.

[673] S. M. Wyatt, "Non-users also matter: The construction of users and non-users of the internet", *Now users matter: The co-construction of users and technology*, pp. 67–79, 2003.

[674] A. Xu, H. Liu, L. Gou, R. Akkiraju, J. Mahmud, V. Sinha, Y. Hu, and M. Qiao, "Predicting perceived brand personality with social media", in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 10, 2016.

[675] C. Yilmaz and E. Ergun, "Organizational culture and firm effectiveness: An examination of relative effects of culture traits and the balanced culture hypothesis in an emerging economy", *Journal of world business*, vol. 43, no. 3, pp. 290–306, 2008.

[676] D. W. Yoo and M. De Choudhury, "Designing dashboard for campus stakeholders to support college student mental health", in *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 2019, pp. 61–70.

[677] A. Zajacova, S. M. Lynch, and T. J. Espenshade, "Self-efficacy, stress, and academic success in college", *Research in higher education*, vol. 46, no. 6, pp. 677–706, 2005.

[678] D. Zapf, C. Dormann, and M. Frese, "Longitudinal studies in organizational stress research: a review of the literature with reference to methodological issues.", *Journal of occupational health psychology*, vol. 1, no. 2, p. 145, 1996.

[679] T. R. Zenger and S. G. Lazzarini, "Compensating for innovation: Do small firms offer high-powered incentives that lure talent and motivate effort?", *Managerial and Decision Economics*, vol. 25, no. 6-7, pp. 329–345, 2004.

[680] D. Y. Zhang, R. Han, D. Wang, and C. Huang, "On robust truth discovery in sparse social media sensing", in *2016 IEEE International Conference on Big Data (Big Data)*, IEEE, 2016, pp. 1076–1081.

[681] D. Zhang, B. Guo, and Z. Yu, "The emergence of social and community intelligence", *Computer*, vol. 44, no. 7, pp. 21–28, 2011.

[682] H. Zhang, M. De Choudhury, and J. Grudin, "Creepy but inevitable?: the evolution of social networking", in *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, ACM, 2014, pp. 368–378.

[683] X. Zhang, L. R. Pina, and J. Fogarty, “Examining unlock journaling with diaries and reminders for in situ self-report in health and wellness”, in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 5658–5664.

[684] X. J. Zhu, “Semi-supervised learning literature survey”, 2005.

[685] J. Zide, B. Elman, and C. Shahani-Denning, “LinkedIn and recruitment: How profiles differ across occupations”, *Employee Relations*, vol. 36, no. 5, pp. 583–604, 2014.

[686] M. Zinkhan, K. Berger, S. Hense, M. Nagel, A. Obst, B. Koch, T. Penzel, I. Fietze, W. Ahrens, P. Young, *et al.*, “Agreement of different methods for assessing sleep characteristics: A comparison of two actigraphs, wrist and hip placement, and self-report with polysomnography”, *Sleep medicine*, vol. 15, no. 9, pp. 1107–1114, 2014.

[687] H. Zou and T. Hastie, “Regularization and variable selection via the elastic net”, *J. Royal Stat. Soc.: Series B*, 2005.

[688] H. Þórarinsdóttir, L. V. Kessing, and M. Faurholt-Jepsen, “Smartphone-based self-assessment of stress in healthy adult individuals: A systematic review”, *Journal of medical Internet research*, vol. 19, no. 2, e41, 2017.